AD-A097 715 ILLINOIS UNIV AT URBANA COMPUTER-BASED EDUCATION RESE--ETC F/G 5/10
THE USE OF INFORMATION FROM WRONG RESPONSES IN MEASURING STUDEN--ETC(U)
FEB 80 M BIRENBAUM, K TATSUOKA N00014-79-C-0752
UNCLASSIFIED CERL-E-12 NL

1 OF 1
AD A
097715

END
DATE
FILMED
5-81
DTIC

Computer-based Education Research Laboratory

University of Illinois Urbana Illinois

THE USE OF INFORMATION FROM WRONG RESPONSES IN MEASURING STUDENTS' ACHIEVEMENT

MENUCHA BIRENBAUM

KIKUMI K. TATSUOKA

This research was sponsored by the Personnel and Training Research Program, Psychological Sciences Division, Office of Naval Research, under Contract No. N00014-79-C-0752. Contract Authority Identification Number NR 150-415.

COMPUTERIZED ADAPTIVE TESTING AND MEASUREMENT RESEARCH REPORT 80-1

81 4 14 003 FEBRUARY 1980

14 CERL-E-12, RR-80-1

SECURITY CLASSIFICATION OF THIS PAGE (~~When Data Entered~~)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Research Report No. 80-1	2. GOVT ACCESSION NO. AD-A097	3. RECIPIENT'S CATALOG NUMBER 715
4. TITLE (and Subtitle) (6) The use of information from wrong responses in measuring students' achievement		5. TYPE OF REPORT & PERIOD COVERED 1 August - 20 November 1979
		6. PERFORMING ORG. REPORT NUMBER CERL Report E-12
7. AUTHOR(s) (10) Menucha Birenbaum & Kikumi Tatsuoka		8. CONTRACT OR GRANT NUMBER(s) (15) N00014-79-C-0752
9. PERFORMING ORGANIZATION NAME AND ADDRESS (16) RR04204 Computer-based Education Research Laboratory University of Illinois, Urbana IL 61801		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (17) 61153N; RR042-04; RR042-04-01; NR154-445
11. CONTROLLING OFFICE NAME AND ADDRESS Personnel and Training Research Programs Office of Naval Research (code 458) Arlington, VA 22217		12. REPORT DATE (11) Feb[redacted]1980
		13. NUMBER OF PAGES (12) 62
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (9) Research rept. 1 Aug-20 Nov 79		15. SECURITY CLASS. (of this report)
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

error analysis, information processing strategies, algorithms, response-patterns, diagnostic adaptive tests, integers operations, achievement tests, dimensionality

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Modern test theory is oriented toward measuring individual performances, regardless of the group or the population which the examinee comes from. In accordance with this trend, there is a noticeable shift from norm-referenced tests to criterion-referenced tests in the area of achievement testing. Latent trait models are becoming very popular due to their desirable properties which enable one to get item-free, population-free measures.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

408130

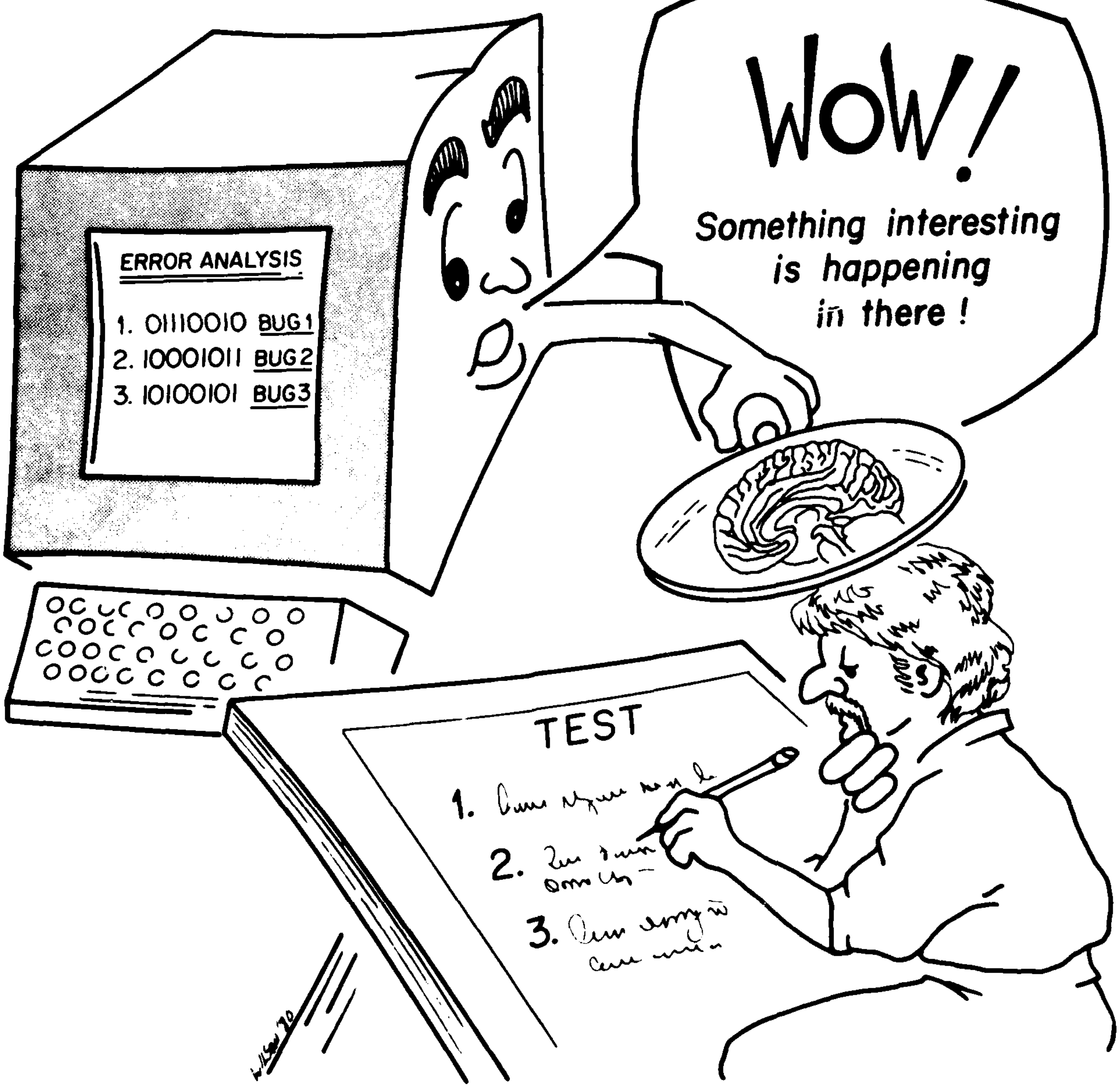
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

However, these models, like those of classical test theory, are using the number of correct responses as the basic source of information for estimating students' achievement.

As good teachers had already realized long ago, a lot of valuable information can be gained by analyzing the students' wrong responses. When a student answers a free response item (s)he gives the response which (s)he considers to be the correct one. Therefore, diagnosing the algorithm that led the student to his/her answer provides an important source of information for asessing his/her achievement.

In light of recent developments in information processing made by cognitive psychologists, as well as by experts in artificial intelligence, it seems that psychometricians ought to "join the party" and incorporate error analsyis as an integral part of their efforts toward improved measurement.

The purpose of this technical report is to provide the reader with some empirical results of error analysis in simple, signed number problems. Types of errors (or wrong algorithms) and their consistency will be defined and discussed. Since a wrong algorithm may sometimes lead to a correct answer the conventional scoring system will be compared with one based on error analysis (i.e., counting a response as correct only if the "correct" algorithm was used) in terms of reliability, latent trait estimates, and the underlying dimensionality.



ERROR ANALYSIS
1. 01110010 BUG1
2. 10001011 BUG2
3. 10100101 BUG3
WOW!
Something interesting is happening in there!
TEST

ACKNOWLEDGEMENT

We wish to acknowledge the kind cooperation extended us by the mathematics teachers at Urbana Junior High School, particularly Ms. Mary Klein, Ms. Virginia Bradford and Ms. Margo Richardson.

Special appreciation is also extended to the students who participated in this research.

Maurice Tatsuoka gave us helpful editorial assistance. Bob Baillie prepared the data collection routines.

Roy Lipschutz & Wayne Wilson provided the excellent art work.

Susan McGrath & Louise Brodie carefully typed the report.

Table of Contents

Introduction.. 1

1. Issues and Methodology.. 3

2. Number Correct--What Does It Tell Us?............................... 7

3. How Stable Are Errors?.. 15

4. A Typology of Alternative Algorithms................................ 20

5. Some Psychometric Properties of a Scoring System
   Based on the Correctness of the Algorithm........................... 27

6. Reaction Time--Is It a Useful Measure of the Strategy
   Underlying the Response?.. 34

Summary and Conclusions... 37

Appendices.. 40

INTRODUCTION

Achievement tests, although customarily treated the same as ability tests by most psychometricians (e.g.,Brown and Weiss, 1977; English, Reckase & Patience, 1977;Bejar and Weiss, 1978;) are designed for different purposes than are most ability tests. Achievement tests are measuring an outcome of a treatment (i.e., instruction). Their main role is to provide the teacher and the student with constructive feedback about the teaching-learning process. Being related to a treatment, achievement tests must satisfy criteria over and aobve those satisfied by most ability tests. In order to enhance the success of the treatment, achievement tests need to have diagnostic capabilities.

When a student responds to a free-response item on an achievement test, (s)he gives the answer which (s)he considers to be the "correct" one. The student's response reflects her/his mental process--the rule (s)he follows in order to reach the answer. As has already been realized by cognitive psychologists, their rules do not always correspond with the algorithm taught in class. As stated by Resnick (1976): "Children seek simplifying procedures that lead them to construct or 'invent' more efficient routines that might be quite difficult to teach directly" (Ibid p. 68). When investigating the relation between the algorithm taught and later performance Resnick concludes that "the efficiency is a result of fewer steps (not, apparently, faster performance of component operations)" and that "the transformation of algorithms by the learner, is more general than we have thought up to now" (Ibid p. 72). Viewing the student's response pattern on an achievement test as reflecting the strategy (s)he uses has important implications for instructional design.

Recent developments in automated diagnostic testing, made by experts in artificial intelligence, also share this point of view regarding achievements tests. The main efforts in this field are directed toward constructing procedural networks that have the capability of identifying student "bugs" or misconceptions. (See, for example: Brown and Burton, 1978).

It should be noticed that considering achievement tests as providing a characteristic response rather than a right or wrong response makes a straightforward analogy to tests in the affective domain, such as personality or attitude tests. This similarity between achievement tests, coming from the cognitive domain, and the tests from the affective domain is apparently due to the common goal those tests serve, as opposed to ability tests, namely the prescription of a treatment. It is where change (improvement or correction) can take place, that we are concerned with an appropriate treatment. We may therefore consider treatment for purposes such as: changing one's self image; changing a person's attitude toward himself, toward other people, objects or issues; improving student achievements; correct student mistakes; etc. It is unlikely, however, to think in terms of treatments in the area of intelligence (e.g., design a treatment for changing one's I.Q.).

It is realized that one may also want to test achievement for purposes other than prescribing further instruction. In situations such as classification or selection one is merely interested in the examinee's level of achievement. However, it may be argued, that no matter what the testing purpose is, achievement tests are measuring an outcome of a treatment. This treatment is aimed at providing the student with the correct algorithm. However, as was already mentioned, the student is most likely to modify that algorithm. This modification can result in a wrong algorithm, which may yield correct answers occasionally, depending on factors such as syntactic attributes of the task presented to the student, that may have led him/her to construct their modified algorithm. Searching for the algorithm reflected in the students' response-patterns may therefore become the gate to more accurate measures of achievement.

Unfortunately, in the current state of psychometric work, this aspect of achievement tests is ignored. Students' responses to an achievement test are judged according to the traditional way as either correct or incorrect. The assessment of achievement is primarily based on the number of correct answers. Even in recent psychometric developments such as adaptive achievement tests, where the main purpose is to tailor the item to the student's level of achievement, no consideration is given to the underlying algorithm or even to the nature of wrong responses (e.g., English, Reckase & Patience, 1977, McKinley and Reckase, 1980).

On the other hand, the cognitive theory of learning and instruction is still far from lending itself to practical applications. As stated by Robert Glaser a few years ago (1976) and which is, unfortunately, true even today, "Experimental psychology of learning and cognition has been almost exclusively a theoretical endeavor, with little effort devoted to application and design of practical techniques for assisting in the conduct of human affairs." On the other hand, as Glaser states: "...psychometrics has become a major technological application of psychology, with primary effort being devoted to practical techniques and less effort to theoretical issues." However, as Glaser emphasizes: "In recent years, there has been increasing interest in and social pressure for the development of professional techniques for the application of what knowledge there is of learning, cognitive processes, and human development. It appears that some linking of theory and practice needs to take place..." (Ibid pp. 1-2).

It seems that error-analysis is one topic in which a joint effort of cognitive psychologists and psychometricians can yield fruitful results in terms of instructional design and measurement. The goal of this technical report is to provide some empirical evidence to support this statement.

ISSUES AND METHODOLOGY

The Issues to be Discussed in This Report:

Based on data collected in seven classes of eighth graders who were studying the topic of signed-numbers, the following issues will be discussed:

1. Can the number of correct answers in a test serve as a sufficient and a meaningful test score for measuring students achievement and for prescribing further instructions?
2. To what extent do students commit stable errors?
3. Can a typology of alternative algorithms, used by students, be identified from analyzing their responses on test items?
4. What is the impact of a scoring system based on the underlying algorithm, as compared to the one based on number correct, on the psychometric properties of the test?
5. Can response time serve as a partial indicator of the mental process underlying the students' responses?

The Data Collection Procedures:

The data to be presented in this report was collected in the fall of 1979 at Urbana Junior High School. Seven classes consisting of 127 eighth graders, taught by two teachers who were using the same instructional method and materials, were observed during the entire instructional period of signed-numbers. The instructional methods, as well as the students' behaviors, were documented by members of our research team.

At the completion of the instruction, a 64-open-ended-item test, consisting of 16 tasks of four parallel items each in addition and subtraction of one or two digit integers, was administered on the PLATO system (A copy of the test is presented in Appendix 1). 127 students took that test and their response patterns, as well as their response times, were stored and analyzed.

The Instructional Method:

Before starting the presentation and discussion of the results, a short outline of the instructional approach will be given since it is conceived of as a crucial component in the identification of the error types.

The instructional unit began by introducing the basic terminology to be used during the teaching process, i.e. "integers," "absolute values," "positive" and "negative" numbers and their location with respect to the number-line. Special notice was given to the number zero. Students practiced solving addition problems using the number line, moving left when adding negative numbers and right when adding positive numbers. In the next stage rules for addition of signed-numbers were given and students were asked to memorize them.

The rules given by the teachers for the addition operation were as follows:

1. For adding two numbers with the same sign, add the absolute values and put the common sign in front of the result;
2. For adding two numbers with different signs follow this two-stage procedure:

Stage 1: Find the difference between the absolute values;
Stage 2: Identify the number with the larger absolute value. The sign of this number will determine the sign of the result.

After students were given drill and practice quizzes and a test in addition of signed-numbers, subtraction was introduced. It was emphasized that any subtraction problem in signed-numbers can be easily converted into an addition problem by following these two steps:

Step 1: Change the operation sign from minus to plus;
Step 2: Change the sign of the second number.

After these two steps are taken, the subtraction problem is converted into an addition problem and should be dealt with according to the rules given for the addition operations.

According to the above-mentioned instructional method the subject matter analysis of signed-number addition and subtraction operations can be schematically expressed as follows:

Addition	A:	+[]++[]	Adding 2 integers with the same sign.
	B:	-[]+-[]	
	C:	-[]++[]	Adding 2 integers with different signs
	D:	+[]+-[]	
Subtraction	E:	+[]--[] → +[]++[] (=A)	Subtraction problems converted into addition problems.
	F:	-[]-+[] → -[]+-[] (=B)	
	G:	-[]--[] → -[]++[] (=C)	
	H:	+[]-+[] → +[]+-[] (=D)	

Table 1 presents the relation of the 64-item test which was administered at the completion of the instructional unit to this subject matter analysis.

Insert Table 1 about here

Figure 1 presents a flow chart for solving signed-number problems according to the insructional method described above.

Insert Figure 1 about here

Table 1

The Subject Matter Analysis and the Corresponding Test Tasks

Task Type Based On the Instructional Method	Test No.	Task Type	Parallel Items No.	
A: +[]++[]	6	L+S	6, 22, 38, 54	A
B: -[]+-[]	10	-L+-S	10, 26, 42, 58	D
	14	-S+-L	14, 30, 46, 62	D I
C: -[]++[]	5	-S+L	5, 21, 37, 53	T
	15	-L+S	15, 31, 47, 63	I O
D: +[]+-[]	3	L+-S	3, 19, 35, 51	N
	11	S+-L	11, 27, 43, 59	
E: +[]--[]	4	S-(-L)	4, 20, 36, 52	
	12	L-(-S)	12, 28, 44, 60	S U
F: -[]-+[]	2	-S-L	2, 18, 34, 50	B
	9	-L-S	9, 25, 41, 57	T
	13	-S-+L	13, 29, 45, 61	R A
G: -[]--[]	1	-S-(-L)	1, 17, 33, 49	C
	8	-L-(-S)	8, 24, 40, 56	T I
H: +[]-+[]	7	L-S	7, 23, 39, 55	O
	16	S-L	16, 32, 48, 64	N

L= Number with larger absolute value
S= Number with smaller absolute value

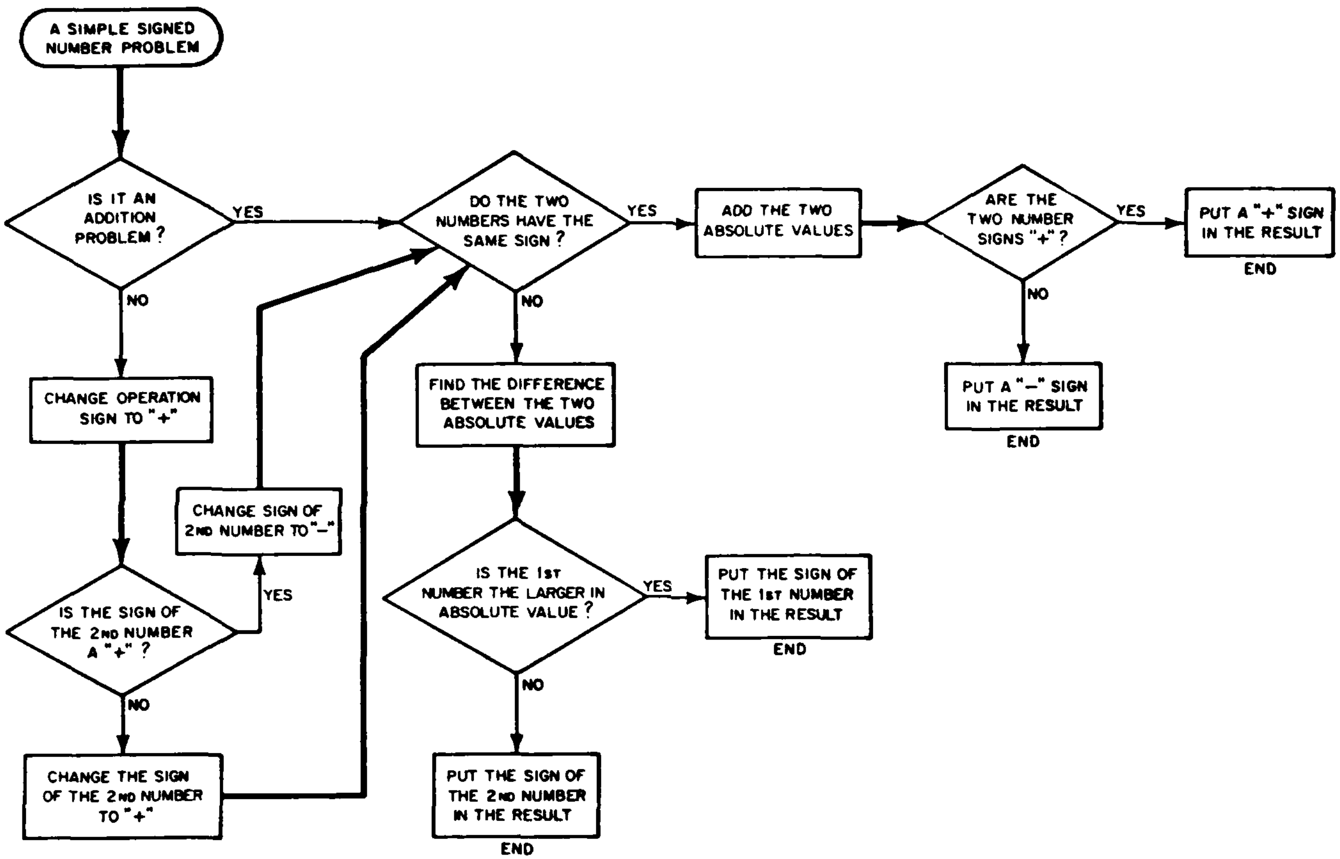


FIGURE 1 : A FLOW CHART FOR SOLVING SIMPLE SI^NED NUMBER PROBLEMS ACCORDING TO THE TEACHING METHOD.

NUMBER CORRECT -- WHAT DOES IT TELL US?

Is the number of correct answers a sufficient and meaningful test score for measuring students achievement and for designing further instructions? As was already mentioned before, most of the testing in the area of achievement, which is done in the classroom is, or at least should be, closely related to the instructional process. Tests are meant to provide the teacher with the necessary information as to how the material is being grasped by the students. Information gained from test answers can therefore serve as a very valuable feedback for the teacher. Unfortunately, most of the teachers, as well as the psychometricians who design and score achievement tests, consider only the correct answers as useful information while totally ignoring information from wrong responses. They treat all kinds of "wrong" responses in the same way, namely, assigning them a score of zero. This scoring system, to say the least, has the effect of pouring out the baby along with the bath water!

Let's consider the following examples presented in Table 2 which are actual responses given by 11 eighth-grade students on a quiz in

<u>Insert Table 2 about here</u>

signed-numbers. (The quiz was given by the classroom teacher at the completion of the instructional unit on addition.) It is obvious from the table that the first three students, even though making the same number of errors, have different kinds of "bugs" or misconceptions concerning the material. The first student treats the parentheses: ( ) as if they were bars representing the absolute value of the number: ||. This misconception of symbols causes him to miss half of the problems in the test. The second student, although scoring the same, shows a procedural "bug" in computing the answer. He consistently misses a step in the process and doesn't distinguish between addition of numbers with the same sign and those of different signs. As was already mentioned before, when the teachers introduced the topic in class, they made a distinction between three kinds of addition problems: when both signs are positive; when both signs are negative; and when each sign is different. For the last case the teachers listed the following two steps to be taken in order to answer the problem:

1. To find the difference between the two numbers;
2. To find the number with the larger absolute value, and to put the sign of that number in the result.

The second student consistently forgot step one. He always added the two numbers and put the sign of the larger absolute value in the result. Using this wrong algorithm he manages to get the correct answer in those cases where the two numbers have the same sign. The third student seems to have no problems with the operatins of signed-numbers. However, the fact that he scored the same as the other two students was due to his

Table 2

Responses of Four Eighth Grade Students on Six Addition Problems in Signed Numbers

Student No. / Problem No.	1	2	3	4
	Responses			
1. 3+−7=−4	−4	−10	−4	10
2. 7+(−3)=4	10	10	4	10
3. −6+−15=−21	−21	−21	−22	−9
4. −6++15=9	9	21	8	21
5. (−23)+(−9)=−32	32	−32	−31	−14
6. (−8)+(−4)=−12	12	−12	−12	−4
No. of Correct Answers	3	3	3	0

weakness in addition and subtraction of whole numbers. Examinations of these three response patterns makes one wonder whether the identical test scores derived by the conventional method of "number correct" indeed tells us anything valuable and useful for the instructinal process or even for estimating the student's level of achievement.

The fourth student, although getting all the problems wrong, has also only one type of "bug" which clearly results from misunderstanding a part of the instruction. When the teachers introduced the number-line, they explained how to move left or right according to the operation sign. This student misunderstood from where he was supposed to start his move for addition. Instead of moving from the location of the first number in the problem, he always moves from the origin (0). When there are two different signs he ends up adding and puts a plus sign in the result. When the two signs are the same he finds the difference and puts the common sign in the result. Does his score of zero on the test indicate the he knows nothing? "Zero knowledge" sounding absurd as it is may also have some very misleading implications in the context of instructional design. Do zero test scores imply that all such students are "tabula rasa," and therefore can be taught the topic by the same method? A positive answer to this question seems profoundly wrong. It assumes that such students get nothing from the former instruction. It fails to recognize that it is the student's interpretation of the former instruction that caused them to develop the algorithm they have been using which, due to some misunderstanding, happened to be wrong.

It clearly follows from the discussion so far, what would be the most appropriate type of instruction to be provided to the students once their "bug" has been identified. It should be an adaptive kind of instruction specifically designed to deal with each student's misconception. It should explain to the student what (s)he is doing wrong and what is the correct way to go. As to the identification of the students' "bugs," a short and efficient procedure should make use of adaptive testing -- adaptive in the sense that each successive problem presented to the examinee should be expected to maximize the amount of relevant information for detecting the algorithm he is using.

In order to carry out a successful adaptive system of testing and instruction, a typology of wrong algorithms needs to be generated. Recent work by Brown and Burton (1978) who developed a procedural network for identifying student "bugs" in subtraction of whole numbers, and by Tatsuoka et al. (1980) who made use of error vectors for identifying error types in signed number operations, provide efficient computerized techniques toward this end.

The discussion so far has focused on the use of achievement tests for the purpose of treatment, i.e. instruction. It is obvious that the use of information from wrong responses in adaptive testing and adaptive instruction have great potential in improving the teaching-learning process. At the same time, error analysis is also valuable in measuring

achievements for purposes other than instruction, e.g., for selection or classification procedures where the main interest is in the achievement level of the examinees. A more concrete and detailed example illustrating the danger of relying on a test score based on the number correct will be discussed in the next section.

Table 3 presents two complete actual response patterns given by eighth graders on the 64-item test of signed-number operations. The two students whose responses are presented in Table 3 received similar total test scores (i.e. number of correct answers). However, a closer examination of their entire pattern of responses reveals a remarkable difference in the kind of misconceptions or "bugs" each of them has. Taking the information provided by error analysis into consideration when judging their achievement level in the topic indicates a large discrepancy between the two in terms of their "true" scores (i.e. the score adjusted for incorrect algorithms).

Insert Table 3 about here

Student 1: The response pattern of this student indicates that he has mastered the addition of signed-numbers. However, he is erring in subtraction problems because of a misconception he has concerning the way of converting a subtraction problem into one in addition. The rule given by the teachers for the conversion was to change two signs: The operation sign and the sign of the second number. The rule student 1 is using through all the subtraction problems is changing one sign only--the operation sign, leaving the other sign unchanged.

The following are examples of the way this student converts subtraction problems into addition. Notice that the fact he has already mastered addition can also be seen in the correct answers he gets on his incorrectly converted problems.

Insert Table 4 about here

Student 2: Although this student got 28 correct answers in the test, a close look at his resonse pattern reveals that most of the correct answers were reached by using an incorrect algorithm. This student knows how to convert a subtraction problem into an addition one. However, he didn't master addition of signed numbers. He hasn't distinguished between adding integers with the same sign and adding integers with different signs. He is always adding the two integers and putting the sign of the larger number in absolute value in the result.

The following table presents some examples of the algorithm used by this student in solving signed-number problems.

Insert Table 5 about here

Table 3

Responses of Two Eighth Graders on Sixty-Four Test Items in Signed Numbers

Problem	Stu. 1	Stu. 2	Problem	Stu. 1	Stu. 2	Problem	Stu. 1	Stu. 2	Problem	Stu. 1	Stu. 2
1. -6-(-8)=2	-14	14	17. -1-(-10)=9	-11	11	33. -3-(-5)=2	-8	8	49. -2-(-11)=9	-13	13
2. -7-9=-16	2	OK	18. -2-11=-13	9	13	34. -4-6=-10	2	10	50. -5-14=-19	9	19
3. 12+(-3)=9	OK	15	19. 7+(-5)=2	OK	12	35. 15+(-6)=9	OK	21	51. 4+(-2)=2	OK	6
4. 1-(-10)=11	-9	OK	20. 3-(-12)=15	-9	OK	36. 5-(-7)=12	-2	OK	52. 6-(-8)=14	-2	OK
5. -3+12=9	OK	15	21. -1+10=9	-9	11	37. -4+13=9	OK	17	53. -2+11=9	OK	13
6. 6+4=10	OK	OK	22. 10+8=18	OK	OK	38. 2+11=13	OK	OK	54. 4+13=17	OK	OK
7. 8-6=2	14	14	23. 7-5=2	12	-12	39. 4-2=2	6	-6	55. 9-7=2	16	16
8. -16-(-7)=-9	-23	-23	24. -12-(-10)=-2	-21	-22	40. -11-(-2)=-9	-13	-13	56. -7-(-5)=-2	-12	-12
9. -12-3=-15	-9	OK	25. -6-4=-10	-2	OK	41. -13-4=-17	-9	OK	57. -9-7=-16	-2	OK
10. -14+(-5)=-19	OK	OK	26. 10+(-1)=-11	OK	OK	42. -7+(-5)=-12	OK	OK	58. -10+(-8)=-18	OK	OK
11. 3+(-5)=-2	OK	-8	27. 2+(-11)=-9	OK	-13	43. 6+(-8)=-2	OK	-14	59. 1+(-10)=-9	OK	-11
12. 9-(-7)=16	2	OK	28. 6-(-4)=10	2	OK	44. 10-(-1)=11	9	OK	60. 13-(-4)=	9	-17
13. -3-(+12)=-15	9	OK	29. -2-(+11)=-13	9	OK	45. -7-(+9)=-16	2	OK	61. -4+(-6)=-10	2	OK
14. -5+(-7)=-12	OK	OK	30. -6+(-8)=-14	OK	OK	46. -2+(-11)=-13	OK	OK	62. -3+(-12)=-15	OK	OK
15. -6+4=-2	OK	-10	31. -5+3=-2	OK	-8	47. -4+2=-2	OK	-6	63. -8+6=-2	OK	-14
16. 2-11=-9	13	-13	32. 5-14=-9	19	-19	48. 7-16=-9	23	-23	64. 4-13=-9	17	-17

Student Number 1: Total OK = 27
Student Number 2: Total OK = 28

Table 4

Example of the Conversion Rule Used by Student Number 1

The Original Subtraction Problem		The Correct Conversion		The Correct Answer	The Student's Conversion		The Student's Answer
-6-(-8)	=	-6+(+8)	=	2	-6+(-8)	=	-14
-7-9	=	-7+(-9)	=	-16	-7+(+9)	=	2
1-(-10)	=	1+(+10)	=	11	1+(-10)	=	-9
8-6	=	8+(-6)	=	2	8+(+6)	=	14
-16-(-7)	=	-16+(+7)	=	-9	-16+(-7)	=	-23
-12-3	=	-12+(-3)	=	-15	-12+(+3)	=	-9
9-(-7)	=	9+(+7)	=	16	9+(-7)	=	2
-3-+12	=	-3+(-12)	=	-15	-3+(+12)	=	9
2-11	=	2+(-11)	=	-9	2+(+11)	=	13

Table 5

Examples of the Algorithm Used by Student Number 2 in Solving Signed Number Problems

Task Number	The Original Problem		The Correct Answer	The Student's Algorithm	The Student's Answer	Student's Answers For the 4 Items In the Task	
						Wrong	Correct
1	-6-(-8)	=	2	-6++8	14	4	-
3	12+-3	=	9	12+-3	15	4	-
5	-3+12	=	9	-13+12	15	4	-
7	8-6	=	2	8+-6	14	4	-
8	-16-(-7)	=	-9	-16++7	-23	4	-
11	3+-5	=	-2	3+-5	-8	4	-
15	-6+4	=	-2	-6+4	-10	4	-
16	2-11	=	-9	2+-11	-13	4	-
2	-7-9	=	-16	-7+-9	-16	3	1
4	1-(-10)	=	11	1++10	11	-	4
6	6+4	=	10	6+4	10	-	4
9	-12-3	=	-15	-12+-3	-15	-	4
10	-14+-5	=	-19	-14+-5	-19	-	4
12	9-(-7)	=	16	9++7	16	1	3
13	-3-+12	=	-15	-3+-12	-15	-	4
14	-5+-7	=	-12	-5+-7	-12	-	4

It is clear from this response pattern that student 2 is still far from reaching mastery in signed-number operations. However, he managed to get almost the same score on this as did student 1 since his wrong algorithm happened to yield the correct answers in the case of the addition problems where the two signs were the same. In task 2 he got only the first item correct; later on, he failed to envision the minus sign in front of the second number, which in this task is the larger absolute value. This explains why he got the answers to the other three problems in this task correct in the absolute value but wrong in the sign.

In a similar task (9) in which the only difference is that the first number rather than the second has the larger absolute value, he managed to get all four items correct. This further supports our interpretations of his incorrect algorithm. It seems clear from the two complete response patterns, which were discussed in detail above, that a valid measure of achievement needs to consider information from wrong as well as from correct responses. The number of correct responses itself may be a misleading indicator since a correct answer can sometimes result from a wrong algorithm. Therefore, in measuring achievement, for any purpose, one would be more accurate if one considers the answers correct only if the algorithm used is the correct one. Information about the correctness of the algorithm can be gained by analyzing the entire response pattern, paying special attention to the error types.

Our discussion so far illustrates the usefulness of error analysis for both measurement and instructional design. Scores based on number of correct answers were shown to be of no diagnostic value at all. Moreover, they can even be misleading in judging the student's achievement level (for some more empirical evidence in this matter, see section 4).

HOW STABLE ARE ERRORS?

The responses to the 64-item test described in the first section were used for evaluating the stability of the errors across parallel items. The different types of responses to the test items were coded according to the following system:

For subtraction problems--

1. The code 1 was given to the correct answer.
2. The code 2 was given when the operation sign was changed but the sign of the second number remained unchanged.
3. The code 3 was given when the operaion sign, as well as the signs of both numbers, were changed.
4. The code 4 was given when the operation sign and the sign of the first number were changed.
5. The code 0 was given for all other computational errors.(1)

For addition problems--

1. The code 1 was given to the correct answer.
2. The code 2 was given when the sign of the second number was changed.
3. The code 3 was given when the sign of the first number was changed.
4. The code 4 was given when the signs of both numbers were changed.
5. The code 0 was given for all other computational errors.(1)

Consistency or stability was defined as getting the same code on at least three out of four parallel items. Table 6 presents frequencies of consistency across the sixteen tasks in the test.

Insert Table 6 about here

As can be seen in the table, less than 6% of the students got inconsistent responses for more than 20% of the tasks. This result indicates a very high proportion of consistent responses. Table 7 presents the frequencies of coded responses for each task separately. Table 8 presents a comparison between frequencies of consistent and inconsistent wrong responses for each task. As can be seen

Insert Tables 7 & 8 about here

from these tables, over 90% of the students mastered the addition tasks. Among those who didn't master addition there is a slight tendency of committing inconsistent errors rather than consistent ones. However, among the subtraction problems, where mastery ranges from 58% to 86%, there are, on the average, more than twice as many consistent errors as there are inconsistent ones.

Table 6

Absolute Frequencies and Percents of Consistent Responses for the Sixteen Test Tasks

Number of Students	Percent	Cumulative Percents	Number of Tasks Consistently Answered
56*	44.8	44.8	16
32	25.6	70.4	15
17	13.6	84.0	14
13	10.4	94.4	13
3	2.4	96.8	12
1	0.8	97.6	11
1	0.8	98.4	10
0	0.0	98.4	9
2	1.6	100.0	8
125		100.0	

*Forty-three of the students answered correctly 15 tasks.

Table 7

Absolute Frequencies and Percentages of Coded Responses for the Sixteen Test Tasks (N = 125)

Task No.	Task Type	Consistent Response Code: Correct 1		E2*		E3		E4		Inconsistent Response	
		N	%	N	%	N	%	N	%	N	%
4	S-(-L)	81	64.8	24	19.2	6	4.8	1	0.8	13	10.4
12	L-(-S)	85	68.0	24	19.2	2	1.6	1	0.8	13	10.4
2	-S-L	73	58.4	27	21.6	2	1.6	2	1.6	21	16.8
9	-L-S	72	57.6	33	26.4	-	-	-	-	20	16.0
13	-S-+L	90	72.0	26	20.8	1	0.8	1	0.8	7	5.6
1	-S-(-L)	93	74.4	9	7.2	5	4.0	8	6.4	10	8.0
8	-L-(-S)	103	82.4	13	10.4	-	-	2	1.6	7	5.6
7	L-S	107	85.6	7	5.6	-	-	1	0.8	10	8.0
16	S-L	88	70.4	4	3.2	2	1.6	19	15.2	12	9.6
6	L+S	124	99.2	-	-	-	-	-	-	1	0.8
10	-L+-S	114	91.2	4	3.2	-	-	1	0.8	6	4.8
14	-S+-L	117	93.6	1	0.8	3	2.4	1	0.8	3	2.4
5	-S+L	115	92.0	1	0.8	3	2.4	1	0.8	5	4.0
15	-L+S	118	94.4	4	3.2	-	-	-	-	3	2.4
3	L+-S	115	92.0	3	2.4	-	-	1	0.8	6	4.8
11	S+-L	115	92.0	1	0.8	3	2.4	-	-	6	4.8

S = Number with smaller absolute value
L = Number with larger absolute value

Error Types:

1. Error Codes for Subtraction:
   a. E2 = Correct Only Operation of Sign
   b. E3 = Changes Operation Sign and Signs of Both Numbers
   c. E4 = Changes Operation Sign and Sign of First Number
2. Error Codes for Addition
   a. E2 = Changes Sign of Second Number
   b. E3 = Changes Sign of First Number
   c. E4 = Changes Sign of Both Numbers

Table 8

Absolute Frequencies and Percentages for Consistent and Inconsistent Wrong Responses (N=125)

Task Number	Task Type	Number of Wrong Responses	Consistent Wrong Responses N	Consistent Wrong Responses %	Inconsistent Wrong Responses N	Inconsistent Wrong Responses %
4	S-(-L)	44	31	70.5	13	29.5
12	L-(-S)	40	27	67.5	13	32.5
2	-S-L	52	31	59.6	21	40.4
9	-L-S	53	33	62.3	20	37.7
13	-S-+L	35	28	80.0	7	20.0
1	-S-(-L)	32	22	68.7	10	31.3
8	-L-(-S)	22	15	68.2	7	31.8
7	L-S	18	8	44.4	10	55.6
16	S-L	37	25	67.6	12	32.4
6	L+S	1	0	0.0	1	100.0
10	-L+-S	11	5	45.5	6	54.5
14	-S+-L	8	5	62.5	3	37.5
5	-S+L	10	5	50.0	5	50.0
15	-L+S	7	4	57.1	3	42.9
3	L+-S	10	4	40.0	6	60.0
11	S+-L	10	4	40.0	6	60.0

S = Number with smaller absolute value
L = Number with larger absolute value

The tasks in which the most consistent errors occur are: Task 1: -L-(-H); Task 4: L-(-H); Task 8: -H-(-L); Task 12: H-(L); Task 13: -L-+H; Task 16: L-H. The most frequent error in those tasks is the one coded as 2 (i.e., forgetting to change the sign of the second number).

An additional finding that can be seen in the presented tables above is that similar tasks, which differ only with respect to the location of the larger number, result in different percentages of correct responses (compare tasks 7 and 16; 1 and 8; 10 and 14; 5 and 15). Some of them differ also in the percentage of consistent wrong responses. Moreover, notations such as brackets or explicitly writing a plus sign in front of the second number do make a lot of difference, as can be seen from the students' responses. (Compare for example tasks 13 and 2; 13 and 1 in Table 8.)

Such results couldn't be predicted on the basis of the instructional method and its underlying subject matter analysis. These results, therefore, support our conclusion that some students get correct answers by applying wrong rules, or, stated another way, some students are using alternative algorithms which result from their misinterpretation of the algorithms introduced in the instructional process. These alternative algorithms occasionally happen to yield the correct answer and thus may mislead us in trying to understand the student's "bug." We therefore suggest that the entire response pattern of the student be analyzed in order to identify his/her algorithm rather than looking only at his/her wrong answers.

A TYPOLOGY OF ALTERNATIVE ALGORITHMS

In the search for student "bugs" that were evident from the high consistency level of errors, the entire response pattern of each student who committed stable errors was carefully analyzed. This analysis, which took the form of a search for the most efficient rule that explains the specific response pattern, was done by three judges independently.

The algorithm identified from the students' response pattern represents the most efficient rule we have found that can reproduce those patterns. We do not claim that these algorithms indeed represent the actual cognitive process that was going on in the students' head. Such a claim would have been too pretentious at this point since we do not have enough information in order to be able to offer a psychological validation. One way of validating the cognitive process is to conduct clinical interviews. However, one should take the conclusions based on such interviews with a grain of salt. In interpreting some protocols from clinical interviews, it was found that the interviewer's responses (verbal or nonverbal) even when not intended to be judgemental of the student's performance, sometims cause a pattern which couldn't be explained otherwise (Resnick 1980).[2] Another way to validate the cognitive process would be to measure response-time. (A short discussion of this topic will be presented in section 6).Since we used neither of those methods extensively in this research study we consider the algorithms we have identified as speculations of the student's cognitive process and suggest taking them just as approximations to be used for adaptive instruction purposes. We do believe that by using the algorithmic approach one can improve the efficiency of the instruction by "debugging" the misconception refelcted in the identified algorithm.

It should be understood that by no means do we suggest, at this point to adjust students' responses on the basis of the presumed correctness of their algorithm, for purposes other than diagnosis or research. We are aware of the unfairness which may be involved in such adjustments, as long as we are not certain that the identified algorithm is indeed the one actually used by the student in deriving his/her answer. (So far, we have been adjusting students scores for research purposes only, the results of that part of the study are reported in Section 5 of this report). We fully agree with the opinion presented by Cazden (1976) in this matter. In his article entitled "On the implication for instructional research", Cazden states that: "...it is essential to remember that a formal analysis of some knowledge or skill doesn't necessarily, or even probably, reflect the organization in anyone's head, much less how it got there.... If we intend only efffective instruction, then the justifying criterion of effectiveness may be sufficient.... But if a model of cognitive process is sought, then more thorough psychological validation is required. Anyone engaged in such endeavors should read Holt's satire of task analysis in his description of how children would be taught, presumably less

sucessfully, how to talk, ... if only to be sure where and how he is wrong" (1976, p. 321).

Bearing this in mind we hereby present a short description of the criterion and the process we have been using for identifying the algorithms. In the topic of signed-numbers there may sometimes be more than one way of explaining the rules of operation beyond a given response pattern. In deciding upon the nature of the algorithm we considered the following criteria:

1. The rule underlying the response pattern should be as closely related as possible to what has been taught in class;
2. The rule should explain the response pattern in the most parsimonious way;
3. The whole response pattern should be reproduced by that rule; (allowing exceptions with regard to tasks 7 and 16 which may have been perceived by some students as belonging to a different category or schemata, i.e., that of "whole numbers", in which the "take away" notion is most commonly being used).[(3)]

The process of identifying the algorithm was essentially one of the following mental hypothesis testing routine. We have hypothesized a priori some misconceptions or "bugs" that may occur as a result of the instructions, those can be described as "discharges" along the network presented as a flow-chart in Figure 1. Those hypothetical "bugs" were used to generate response patterns for the 16 test tasks. The raw answers were translated into the codes as described in section 3 and these were matched with the actual resonse patterns given by the students. We realize that since the instructional method emphasized the rules for solving signed-number problems, some students could have correctly answered the test by memorizing those rules without having an insight or full understanding as to why those rules are applied. We are not intending to judge the algorithms at that level since a correct response pattern provides no additional information as to the level of understanding. In this report we refer only to those records that include some incorrect responses (which are consistent across parallel items).

Appendix 3 presents a detailed list of those algorithms including the codes given to the answers according to a coding system, and the specific response which uniquely identifies each alternative algorithm. The following is a summary list of the algorithms identified in our current data:

1. Always subtracts and puts the sign which appears in front of the larger number in absolute value. As can be seen in the table, consistently following this rule results in correct answers for seven out of the sixteen tasks. However, it implies failure to distinguish between addition and subtraction operatins, as well as between operation-signs and number-signs.

2. In addition problems--Always adds and puts the sign of the number with the larger absolute value.

In subtraction problems--Converts correctly into addition (according to the method taught in class) but fails to carry out the addition because of the aforementioned "bug". Nevertheless, the student using this algorithm manages to get half the test tasks correct, as can be seen in the table.

3. In the addition of signed numbers, the student finds the difference between the two numbers and assigns the sign of the number with the larger absolute value to the result. In subtraction (s)he follows these three steps: First adds the absolute values, then changes the sign of the operation and that of the second number, and finally assigns the result the sign of the number with the larger absolute value. Following this strategy the student manages to get correct answers for 10 test tasks.

4. Whenever the problem consists of two similar signs the student adds and puts a plus sign in the result. For problems with two different signs, the student subtracts and puts a minus sign in the result. Thus, without being able to distinguish between addition and subtraction problems, the student manages to get correct answers for three test tasks.

5. Using the number line to figure out his jumps, the student mistakenly always jumps from the origin (0) instead of jumping from the value of the first number in the problem. This "bug" causes the student to miss all the test items, even though (s)he is able to differentiate between the operations and between the number signs.

6. In subtraction, the student always forgets to change the sign of the subtrahend. Obviously, this "bug" causes him/her to miss all the subtraction tasks.

7. In subtraction, the student doesn't change the signs of the operation and the subtrahend. (S)he always treats the parentheses as if they were bars indicating the absolute value of the number. This combination of "bugs" results in correct answers for three out of the nine subtraction problems. (Notice that, even though this student shows two "bugs" in the subtraction problems, his/her score based on the number correct is higher than the score of a student following algorithm 6 who has a relatively minor misconception).

8. In subtraction, the student puts a plus sign in front of the subtrahend (i.e. changes a minus to a plus but doesn't change a plus to a minus). Tasks 7 and 16 seem

to "belong" to another "box" because of the missing signs. Once the conversion has been carried out, the student proceeds correctly with the converted addition problems.

9. In subtraction, the student always subtracts and puts the sign of the number with the larger absolute value.

10. In subtraction, the student changes only the operation sign into addition when there is a sign in front of the subtrahend. When the plus sign of the subtrahend is missing, the student considers the missing sign as the operation sign. Following this algorithm (s)he manages to get correct answers for four out of the nine subtraction problems in the test. Once the conversion is done (s)he proceeds solving correctly the converted addition problems.

11. In subtraction, the student changes the signs of both numbers as well as the operation sign. This wrong conversion causes him/her to miss all the subtraction problems in the test.

12. The student treats parentheses as if they were bars indicating the absolute value. Since parentheses were included only in the subtraction problem, this minor misconception of symbols results in incorrect answers to four out of the nine subtraction tasks in the test.

13. In addition, the student changes the sign of the second number (applying incorrectly the rule for subtraction to the addition problem). Following this rule (s)he manages to get all the addition problems incorrect while demonstrating mastery in solving the subtraction problems.

14. In subtraction, the student subtracts the absolute values and attaches the sign of the larger absolute number to the results (except for problems with three minus signs, i.e. Tasks 1 and 8).

In addition, the student adds the absolute values and assigns the sign of the number with the larger absolute value to the result, except for Task 3 which (s)he perhaps approaches in an intuitive manner using the "take away" idea. Following this strategy, the student manages to get correct answers for seven tasks on the test.

15. In subtraction, the student subtracts the absolute values and assigns the result the sign of the number with the larger absolute value, except for the case where there are three minuses (i.e. Tasks 1 and 8) where (s)he adds and attaches the common sign to the result.

In addition, the student subtracts and attaches the sign of the number with the larger absolute value to the result

except for Task 6 which doesn't involve minus signs. Thus, being unable to distinguish between addition and subtraction of signed-numbers, the student manages to get correct answers for six test tasks.

16. In subtraction, the student subtracts the absolute values and attaches to the result that sign which appears in front of the number with the larger absolute value. However, when the problem involves three minuses, the rule is modified. The student still subtracts absolute values, but assigns a plus sign to the result. Thus, being unable to distinguish between operation signs and number signs, the student manages to get correct answers for three out of the nine test tasks.

It is clear from looking at those alternative algorithms that some students are using a single rule which they apply to all types of problems regardless of the operation signs. Others, however, have invented a more complicated algorithm which is determined by factors such as the location of the number with the larger absolute value in the problem or whether or not the sign is explicitly written in the problem. A subtraction problem with three minuses turns out to one that causes a "modification" of the rule in some cases. However, addition problems where the two numbers are positive and their signs are missing were answered correctly by almost all the students (99%), regardless of the rule they have been using for solving other problems. In subtraction problems of that kind, the ones where the first number was the larger in absolute value were answered correctly by 86% of the students, again regardless of the rule they have been using for solving other subtraction problems (including a similar task where the number with the smaller absolute value came first). It is pretty obvious that these problems were not perceived as "signed-number operations" and students were capitalizing in their response on the "take away" idea they had been taught for subtraction of whole numbers in earlier grades. In the current data the most common of the alternative algorithms described above are 6, 8 and 9. It is clear that some alternative algorithms result from a single "bug" whereas others arise from a combination of "bugs." Those vary in their seriousness and the number of steps to be taken for correction. It is evident from the current data that while addition was mastered pretty well by the great majority of students, the subtraction procedure caused some confusion. However, several students failed to distinguish between addition and subtraction operations. Others failed to distinguish between operation signs and number signs. Some students applied the subtraction rules in solving addition problems whereas others had troubles with the notations.

It is clear that these differences in student "bugs" or misconceptions could not have been detected without analyzing the complete response pattern including incorrect as well as correct responses.

Figure 2 presents a flow-chart for solving simple signed-number

problems. This flow chart is based on the information gained from

Insert Figure 2 about here

analyzing the students response patterns with respect to the underlying algorithm.

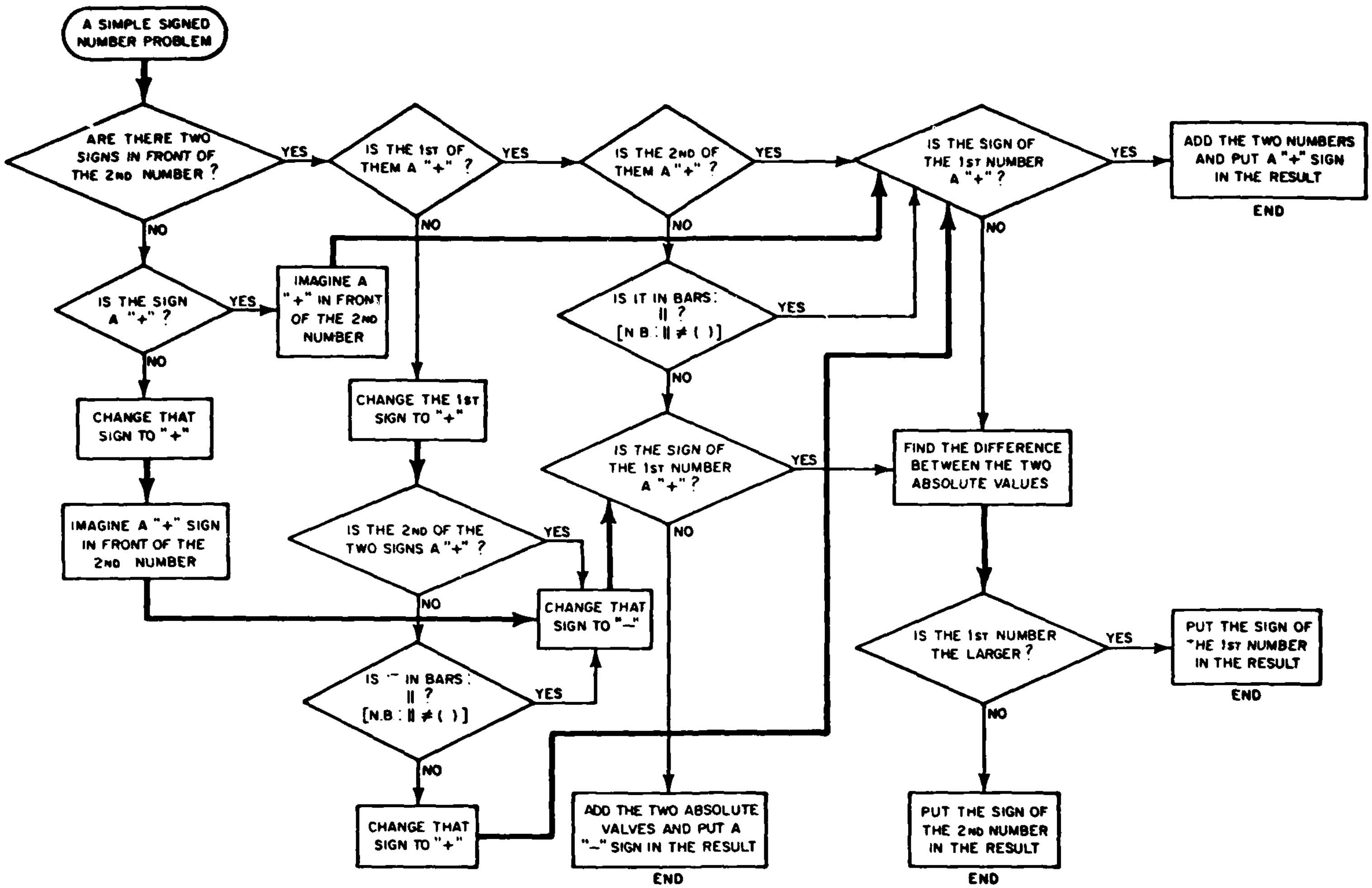


Figure 2 : A FLOW CHART FOR SOLVING A SIMPLE SIGNED NUMBER PROBLEM (CONSTRUCTED ON THE BASIS OF THE ERROR ANALYSIS)

SOME PSYCHOMETRIC PROPERTIES OF A SCORING SYSTEM BASED ON THE CORRECTNESS OF THE ALGORITHM

Since some wrong algorithms happen to yield correct answers, as was illustrated in the previous chapter, a scoring system based on the presumed correctness of the algorithm was adopted. Each consistent response pattern was carefully analyzed and its underlying algorithm (the rule which most efficiently reproduced such a pattern) was identified. Only right answers that were determined by the correct algorithm were credited; all other correct responses were assigned a score of zero (including correct answers that resulted from wrong algorithms). Applying this scoring system resulted in adjusting 32 out of the 125 students' records. Figure 3 shows the number of scores adjusted in each task.

Insert Figure 3 about here

In order to test the effect of the scoring system, based on the error analysis results, on the psychometric properties of the test, a comparison between this method and the conventional one in terms of reliability, dimensionality and latent trait estimates was carried out.

Procedure:

Following the conventional scoring system, a score of 1 was given to the correct answer and 0 otherwise.

For the scoring procedure based on the error analysis results, a score of 1 was given only when the correct answer was presumed to be derived by the correct algorithm. Otherwise a score of 0 was assigned to the answer. Since the test consisted of four parallel items for each of the sixteen tasks, a task score was computed for each of the two scoring systems. A " mastery task score" of 1 was assigned to a student who got a correct answer on at least three out of the four parallel items. Otherwise a task score of 0 was assigned. The student's response patterns consisting of two binary vectors (corresponding to the two scoring systems) for the 16 test tasks were compared in terms of the correlation matrices among the tasks, the reliability coefficients (α's), the eigenvalues, and the oblique factor pattern matrices. The computer package SPSS (Nie et al., 1975) was used for those analyses. A computer program for multidimensional scaling KYST (Kruskal et al., 1973) was used for mapping the task scores as points in a two-dimensional space. The two latent-trait parameters (a's and b's), the discrimination and difficulty indices were computed via the "get ab" program implemented on the PLATO system by R. Baillie.

Results:

A. Reliability

The standardized item reliability coefficient α (Cronbach

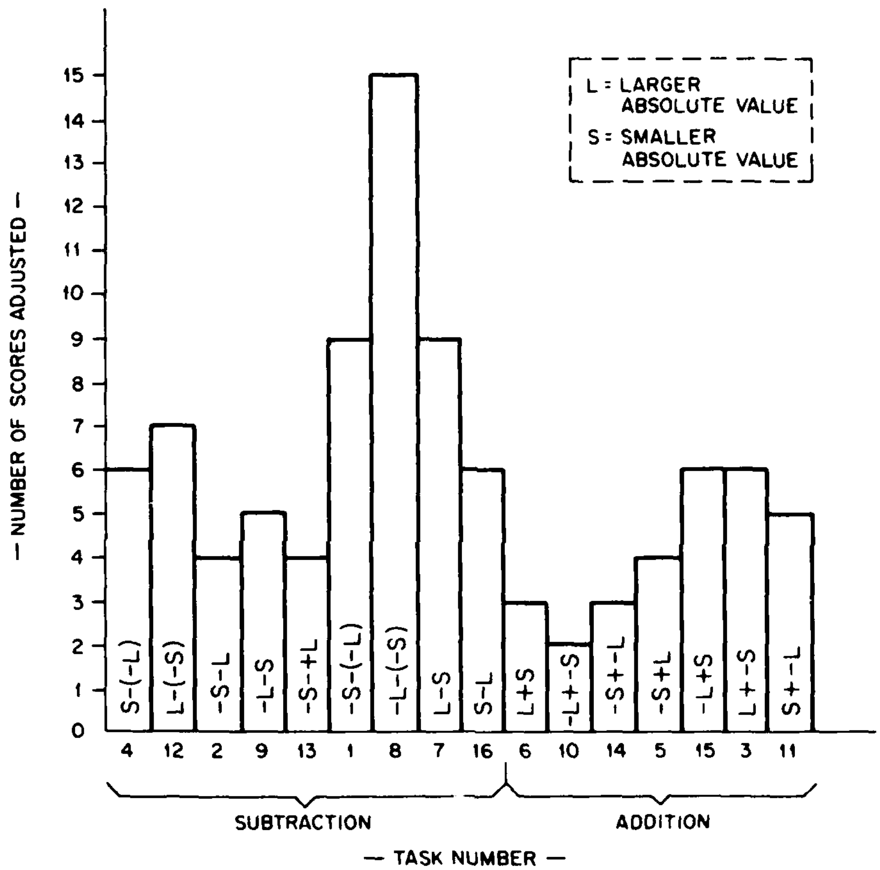


Figure 3

NUMBER OF SCORES ADJUSTED IN EACH TEST TASK

1951) on the 16 task scores for 125 subjects was .84 for the unadjusted rescoring (based on the conventional scoring procedure) and .93 for the adjusted rescoring based on the underlying algorithm.

B. Dimensionability

The correlations among the test task scores before and after adjusting is shown in Appendix 4. As can be seen in the table, there is a drastic change in the correlation matrix after the adjustments were made. There are no negative correlations among the tasks as there were before the rescoring. Ninety-nine out of the total of 120 comparisons between pairs before and after adjusting the scores yield significant differences.[4] The reduction in the dimensionality of the adjusted scores is evident from the comparison of the eigenvalues derived from a principal component analysis. Those eigenvalues and the percent of variance explained by them in the two sets of scores (before and after adjustments) are presented in Table 9 and in Figure 4. As can be seen in the table, only two eigenvalues exceed unity in the adjusted scores as compared to four in the original data.

Insert Table 9 about here

Morever, the two eigenvalues account for a larger percent of variance than do the four eigenvalues (77.9% in the two-eigenvalue case as compared to 71.9% in the 4-eigenvalue case).

Insert Figure 4 about here

A two-dimensional plot of the test task before after the adjustments is presented in Figure 5. As can be seen in the figures, there is a remarkable difference in the configuration of the te _ points in the plan before and after the adjustment. While before the adjusting, the plot looks chaotic, after rescoring, two distinct sets of points are clearly identifable, one consisting of the subtraction problems and the other of the addition ones. An oblique factor

Insert Figure 5 about here

analysis using the square multiple correlations as initial estimates of the communalities distinctly shows the pattern of the test task loadings on the two factors. Appendix 5 presents this pattern matrix. As can be seen in the table, the first factor is highly loaded with all the subtraction problems and the second one with all the addition problems.

C. Latent-Trait Estimates:

The effect of the scoring system based on the presumed correctness

Table 9

Eigenvalues and Percent of Explained Variance of Test Task Scores Before and After Adjustment (N=125)

λ	Before adjusting % Of Variance	λ	After adjusting % Of Variance
6.289	39.3	9.361	58.5
2.632	16.5	3.109	19.4
1.511	9.4	.751	4.7
1.080	6.7	.485	3.0
.989	6.2	.397	2.5
.573	3.6	.327	2.0
.552	3.4	.284	1.8
.434	2.7	.263	1.6
.394	2.5	.221	1.4
.350	2.2	.190	1.2
.314	2.0	.153	1.0
.264	1.7	.136	.9
.207	1.3	.123	.8
.167	1.0	.085	.5
.124	.8	.069	.4
.119	.7	.045	.3

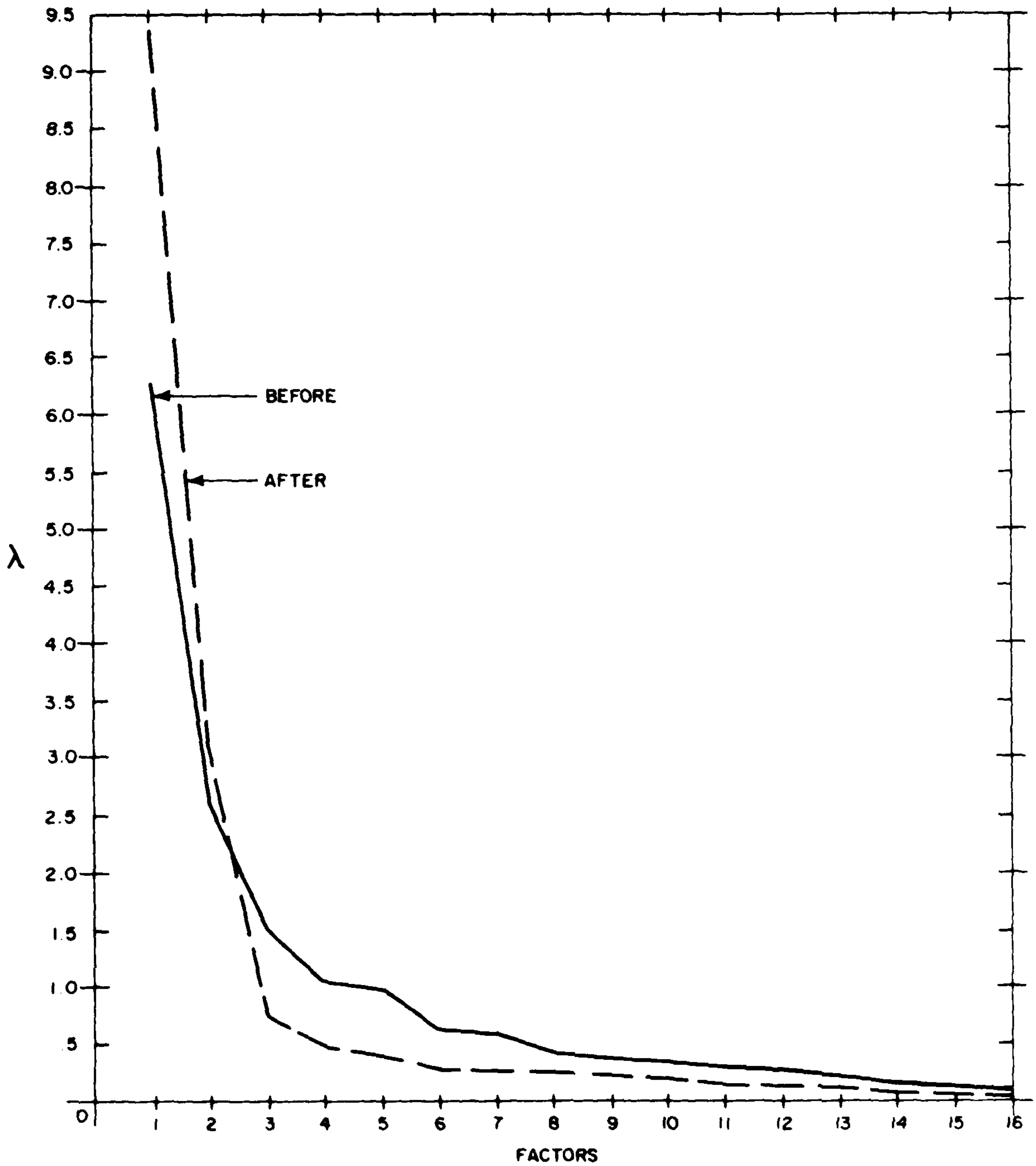


Figure 4 Screetest - Eigenvalues extracted in a principal component analysis before and after adjusting the scores.

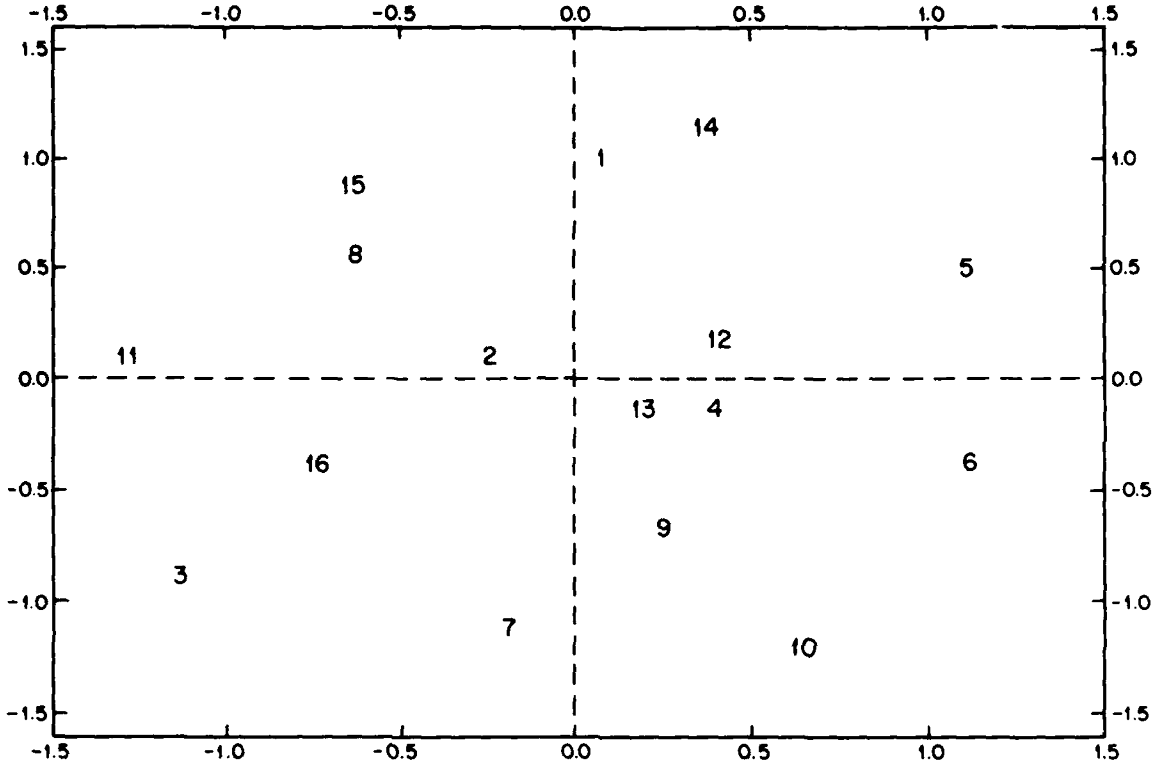


Figure 5a: ORIGINAL TASK SCORES
A 2 DIMENSIONAL PLOT

TASK NO.	TASK TYPE
1 =	-S-(L)
2 =	-S-L
3 =	L+-S
4 =	S-(-L)
5 =	-S+L
6 =	L+S
7 =	L-S
8 =	-L-(-S)
9 =	-L-S
10 =	-L+-S
11 =	S+-L
12 =	L-(S)
13 =	-S-+L
14 =	-S+-L
15 =	-L+S
16 =	S-L

L = LARGER ABSOLUTE VALUE
S = SMALLER ABSOLUTE VALUE

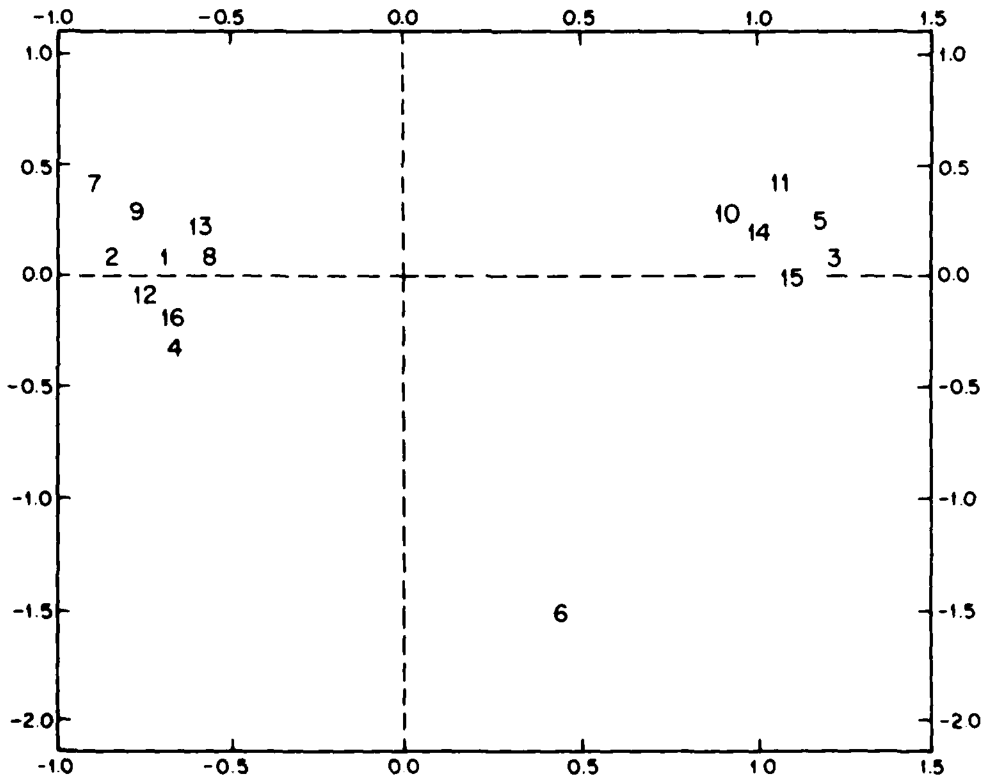


Figure 5b: ADJUSTED TASK SCORES
A 2 DIMENSIONAL PLOT

of the algorithm as compared to the conventional scoring system was studied in respect to latent-trait estimates. The two-parameter logistic model for estimating the difficulty (b's) and discriminating power (a's) indices was applied to the set of subtraction problems before and after adjusting the scores. We did not combine the addition and subtraction problem since the ICC model we used assumes unidimensionality (Lord & Novick 1968). As was shown earlier, the dimensinality of the data, as tested by factor analysis was larger than one in both before and after adjusting the scores. In the latter case, where the dimensionality was crystalized, two dimensions emerged, the first of which contained the subtraction problems. This factor accounted for the greater proportion of variance. The second factor contained the addition problems. Since most of the subjects reached mastery on these problems, the parameter estimation by the maximum likelihood was not carried out. However, when this technique was applied to the set of subtraction problems it took only a few iteration for the estimation procedure to converge for the adjusted data (after deleting task 7: L - S type of problem which doesn't require knowledge in signed numbers). When applied to the same set of items on the original data (before adjusting the scores) the ML procedure did not converge. Thus, we could not estimate the a's and b's for this data-set.

REACTION TIME -- IS IT A USEFUL MEASURE OF THE STRATEGY UNDERLYING THE RESPONSE?

As is evident from the discussion so far, the number of correct answers turns out to be insufficient, and sometimes even a biased measure of student achievements.

We have already demonstrated cases in which the student's tendency to modify the algorithm taught in class results in an incorrect modified rule that occasionally yields correct answers. Therefore, by looking only at the number of correct answers, one cannot get accurate information as to the correctness of the strategy used by the student.

An additional measurable source of information in the student's response is reaction time. Lachman _et_ _al_. (1979) recently stated that "The current state of choice reaction time has moved from a topic of study to a methodological tool. It has a well-developed theoretical framework, and its properties are rather clearly understood. Because it is so useful, it is used in virtually every area of cognitive psychological research." (Ibid. p. 182).

Cognitive psychologists specializing in instruction have realized that when measured accurately, response latencies can provide fairly good indicators as to the number of steps the individual takes in solving a problem (See, for example: Woods et. al, 1975; Groen and Parkman, 1972; Resnick, 1978). Thus, examination of response time may provide valuable information for identifying the strategy used by the student, and in this way contribute to a better dignosis of his/her "bugs" and to improved estimations of his/her achievement level.

In the current study, we have collected response times for the 64-item test, but since the test wasn't a priori designed for this purpose, the a posteriori comparisons that can be made are limited. As was described in Section 1, each task on the test consisted of four items that were matched on operation signs, number signs and the location of the larger absolute number in the problem. However, the absolute numbers, their sums and differences were controlled only with respect to the number of digits and the number of keys to be pressed for the correct answer. As a result, response time on parallel items such as 34 (-4-6=-10) and 50 (-5-14=-19) is not expected to be equal due to the different values of their differences. As has already been demonstrated in response time studies, for a given model of subtraction or addition, latencies rise as a function of the value of the number to be decreased. (See, for example: Woods et. al, 1975; Groen and Pakman, 1972).

As can be seen from our data too, for students who mastered[5] all the 16 test tasks (to be referred to hereafter as "experts"), the mean latency for item 34 is 10.104 sec., whereas for item 50 it is 13.718

sec.[6] The difference between these two means, as tested by a t-test for dependent samples, turns out to be significant ($t=2.16$; $p<.05$). Due to this restriction, averaging latencies across task items seemed illegitimate. Thus, we had to resort to comparisons of the item level. Since those comparisons involved only problems with identical numbers, obviously not all the task latencies could have been compared to one another.

Another constraint to be considered when comparing latencies is the location of the problems in the test. In responding to similar tasks, the response to the second is usually faster than the first. This fact, according to Lachman et. al (1979) can be taken as "evidence that the person is capitalizing on an already activated 'pathway' in making the second reponse." (Ibid. p. 181). Tatsuoka and Tatsuoka (1979) have already demonstrated this phenomenon with a similar test of signed number operations. Table 10 presents the comparisons among item latencies that were made under the above mentioned constraints. Those results are based on data from "experts" only and therefore are not contaminated by latencies for incorrect strategies yielding correct answers, nor are they affected by partial knowledge or guessing.

As can be seen in Table 10, most of the comparisons (71%) indicate significant differences among the compared means. Moreover,

Insert Table 10 about here

except for two comparisons, all the other significant comparisons confirm the a priori hypothesized directionality that was based on the teaching method. These results clearly indicate that the more steps involved in the solution, the more time it takes to arrive at the correct answer. The two exceptions to this generalization involve items from tasks 7 and 16. It has already been mentioned before that these tasks, in which two plus signs are missing, were probably not conceived as related to signed-number operations. Thus, the great majority of students were solving those tasks by using the whole-number subtraction or addition schemata which they have already mastered in earlier grades.

Table 10

Comparisons of Latencies of Different Problems With Identical Numbers

Comparison Number	Problem Type[1]		No.	Hypothesized Directionality	N[2]	$\bar{X}$	S.D.	t[3]	Sig.[4]
1	E	1-(-10)=11	4	E > B	30	12,132	5350	4.45	**
	B	-10+(-1)=-11	26			7,663	2525		
2	E	10-(-1)=11	44	E > B	32	10,345	5211	3.04	**
	B	-10+(-1)=-11	26			7,617	2531		
3	D	1+(-10)=-9	59	D > B	31	9,724	4702	2.30	*
	B	-10+(-1)=-11	26			7,734	2551		
4	H	9-7=2	55	H > E	31	10,396	5209	-1.73	N.S.
	E	9-(-7)=16	12			12,656	5511		
5	H	9-7=2	55	H > F	33	9,967	5243	-1.81	N.S.
	F	-9-7=-16	57			12,232	5379		
6	E	9-(-7)=16	12	F > E	31	12,656	5511	1.78	N.S.
	F	-7-(+9)=-16	45			10,609	4004		
7	E	-6-(-8)=14	52	E > D	34	11,318	4184	2.70	*
	D	6+(-8)=-2	43			8,623	4096		
8	F	-2-11=-13	18	$F_3 > F_1$	28	12,571	5952	3.20	**
	F	-2-(+11)=-13	29			9,722	3472		
9	H	2-11=-9	16	H > A	34	12,938	6048	3.44	**
	A	2+11=13	38			8,481	4156		
10	A	2+11=13	38	B > A	35	8,715	4600	-1.64	N.S.
	B	-2+(-11)=-13	46			11,235	8488		
11	B	-2+(-11)=-13	46	C > C	34	11,348	8589	1.61	N.S.
	C	-2+11=9	53			8,864	3693		
12	G	-2-(-11)=9	49	G = G	32	14,566	7287	1.85	N.S.
	G	-11-(-2)=-9	40			12,040	4498		
13	F	-2-11=-13	18	$F_3 > F_1$	28	12,571	5952	3.20	**
	F	-2-(+11)=-13	29			9,722	3472		
14	C	-2+11=-9	53	D > C	34	8,925	3677	-2.52	*
	D	2+(-11)=-9	27			11,489	5392		
15	A	6+4=10	6	E > A	31	7,560	2807	-3.17	**
	E	6-(-4)=10	28			12,167	8167		
16	H	7-5=2	23	H > E	33	8,752	4685	-3.04	**
	E	5-(-7)=12	36			12,331	7077		
17	C	-5+3=-2	31	G > C	33	8,900	5155	-1.82	N.S.
	G	-3-(-5)=2	33			11,388	6205		
18	D	3+(-5)=-2	11	G > D	30	7,300	2339	-3.76	**
	G	-3-(-5)=2	33			11,109	6283		
19	C	-4+13=9	37	C > A	32	10,879	4163	3.70	**
	A	4+13=17	54			8,017	3268		
20	H	4-13=-9	64	H > A	32	14,299	7889	5.55	**
	A	4+13=17	54			7,371	1848		
21	A	4+13=17	54	C > A	32	8,017	3268	-3.70	**
	C	-4+13=9	37			10,879	4163		
22	E	13-(-4)=17	60	E > A	35	11,701	4180	6.51	**
	A	4+13=17	54			7,774	3163		
23	H	4-13=-9	64	H > E	31	14,202	8000	2.16	*
	E	13-(-4)=17	60			11,350	3895		
24	H	7-5=2	33	H > E	33	8,752	4685	-3.04	**
	E	5-(-7)=12	36			12,331	7077		

(1) The letters represent the hierarchy of the subject matter analysis based on the instructional model as described in Chapter 1 of this report.

(2) This data is based on "experts'" responses (only those students who mastered all the sixteen test tasks).

(3) t-Test for dependent samples.

(4) * = $p < .05$; ** = $p < .01$

SUMMARY AND CONCLUSIONS

Error analysis--Don't it make no-nevermind? The main focus of this report was on issues specifically related to measurement of achievement tests. The differences between achievement tests and ability tests were discussed and empirical data was used to illustrate the paucity of information conveyed by test scores based on the number of correct answers in measuring achievement.

The data used consisted of responses to a 64-item test on signed-numbers, responded to by 127 eighth graders at the completion of an instructional unit in signed numbers. The test consisted of 16 tasks, each of which had four parallel, open-ended items. It was administered on PLATO and reaction-time was stored for each response besides the response itself. Using this data, the stability of errors across parallel items was tested and it was shown to be pretty high for a vast majority of students. A close examination of the entire response pattern for those students who committed consistent errors resulted in identifying a typology of 16 alternative algorithms used by students for solving signed number operations. These algorithms varied with respect to the number and the seriousness of the "bugs" or misconceptions students developed concerning the subject matter. The idea that students tend to modify the algorithm taught in class is already a well recognized one among cognitive psychologists. The fact that some of those modifications result in incorrect algorithms should concern psychometricians as well, especially in the cases where wrong algorithms happen to yield correct answers, as was often the case in our study. Respose time was shown to be a useful tool in helping to identify the underlying algorithm.

Based on these results it seems necessary in measuring achievements to examine the entire response pattern, considering correct as well as incorrect answers in order to be able to infer the underlying algorithm. We have classified error types, coded them in a uniform way, and then used the coded pattern to identify the algorithm. On the basis of the identified algorithm, we have rescored the test so that only right answers presumably derived by an incorrect algorithm received no credit. This procedure resulted in a substantial gain from the psychometric aspect. The adjusted scores turned out to be superior to the original scores in terms of reliability and latent trait estimates. The dimensionality of the adjusted scores was smaller and the structure became much clearer as compared to that of the unadjusted scores.

These results may have some important implications for psychometric work on achievement tests. "Latent trait" has become a very popular concept in the last few years among psychometricians. Sophisticated mathematical models are being developed for estimating ability from responses to test items. Adaptive testing, in which the items are meant to be tailored to the examinee's level of ability, are constructed using the estimates derived from those models. It seems that transferring this approach to achievement tests requires some modifications due to

the different nature of achievement testing from ability testing. Be the purpose of the achievement test diagnosis or be it for estimating achievement level in selection/classification programs, achievement tests are always measuring an outcome of a treatment (i.e., instruction). That treatment was meant to provide the student with an algorithm to be used in solving problems in a specific subject matter area. The student's response on the test items reflect his/her modification of that algorithm. The aim of the testing should therefore be to identify that latent cognitive process in order to quantify the responses in a meaningful and an efficient manner. Adaptive tests in this context should aim at "debugging" the student's misconceptions so as to enable a complete understanding of his/her algorithm or strategy in solving the problem.

FOOTNOTES

(1) Note: Only four error types are of interest as far as the concept of signed numbers is concerned; (i.e., plus or minus the sum or the difference of the two numbers in the problem). Errors resulting from failure to master addition or subtraction of whole numbers were coded as 0.

(2) Personal communication, April 1980.

(3) Support for this comes from protocols of clinical interviews with students solving the same version of signed-number tests. Those protocols were kindly provided to us by Mr. Seth Chaiklin of the Learning & Development Center at the University of Pittsburgh, to whom we are greatly indebted.

(4)The significance of the difference between each pair of correlations was tested as follows:

$$z = \frac{Z_1 - Z_2}{\sqrt{\frac{2}{N - 3} \quad 1 - R_{min}}}$$

where R min. = r before adjusting or r after adjusting, whichever is smaller.

(5) Task mastery was said to be the case in which a student got the correct answer for at least 3 out of 4 parallel items in a task.

(6) Appendix 6 presents the latencies for all items for the group of "experts".

Appendices

1. The Sixty-Four Item Test in Addition and Subtraction of Signed Numbers (Sixteen Tasks, Four Parallel Items Each) 41

2. Frequencies of Coded Responses to Each Item in the Test of Signed Numbers (N=125) 42

3. A Typology of Alternative Algorithms used by students in Solving Signed-Number Problems 43

3a. Wrong Algorithms that Yield Correct Answers 44

4. Correlations, Means and Standard Deviations Before and After Rescoring Items (Based on the Underlying Algorithms) 45

5. Oblique Factor Pattern Matrix for Rescored Task Scores 46

6. Response Time--Experts 47

Appendix 1

The Signed-Number Test

Test Items			
I	II	III	IV
1. -6-(-8)=2	17. -1-(-10)=9	33. -3-(-5)=2	49. -2-(-11)=9
2. -7-9=-16	18. -2-11=-13	34. -4-6=-10	50. -5-14=-19
3. 12+-3=9	19. 7+-5=2	35. 15+-6=9	51. 4+-2=2
4. 1-(-10)=11	20. 3-(-12)=15	36. 5-(-7)=12	52. 6-(-8)=14
5. -3+12=9	21. -1+10=9	37. -4+13=9	53. -2+11=9
6. 6+4=10	22. 10+8=18	38. 2+11=13	54. 4+13=17
7. 8-6=2	23. 7-5=2	39. 4-2=2	55. 9-7=2
8. -16-(-7)=-9	24. -12-(-10)=-2	40. -11-(-2)=-9	56. -7-(-5)=-2
9. -12-3=-15	25. -6-4=-10	41. -13-4=-17	57. -9-7=-16
10. -14+-5=-19	26. -10+-1=-11	42. -7+-5=-12	58. -10+-8=-18
11. 3+-5=-2	27. 2+-11=-9	43. 6+-8=-2	59. 1+-10=-9
12. 9-(-7)=16	28. 6-(-4)=10	44. 10-(-1)=11	60. 13-(-4)=17
13. -3-+12=-15	29. -2-+11=-13	[illegible]. -7-+9=-16	61. -4-+6=-10
14. -5+-7=-12	30. -6+-8=-14	46. -2+-11=-13	62. -3+-12=-15
15. -6+4=-2	31. -5+3=-2	47. -4+2=-2	63. -8+6=-2
16. 2-11=-9	32. 5-14=-9	48. 7-16=-9	64. 4-13=-9

Appendix 2

Frequencies (Percentage) of Classified Responses
to Each Item in the Test of Signed-Numbers
(N=127)

Item No.	Problem Type	Correct f	Correct %	E2* f	E2* %	E3* f	E3* %	E4* f	E4* %	Others %
4.	1-(-10)=11	11	59	-9	24	9	13	-11	2	2
20.	3-(-12)=15	15	66	-9	23	9	6	-15	2	3
36.	5-(-7)=12	12	65	-2	20	2	9	-12	3	3
52.	6-(-8)=14	14	68	-2	21	2	6	-14	2	4
12.	9-(-7)=16	16	61	2	31	-2	4	-16	2	2
28.	6-(-4)=10	10	69	2	24	-2	6	-10	2	-
44.	10-(-1)=11	11	72	9	21	-9	3	-11	2	2
60.	13-(-4)=17	17	72	9	23	-9	1	-17	2	2
2.	-7-9=-16	-16	56	2	31	-2	7	16	2	4
18.	-2-11=-13	-13	61	9	28	-9	6	13	3	2
34.	-4-6=-10	-10	62	2	30	-2	3	10	3	2
50.	-5-14=-19	-19	63	9	31	-9	2	19	2	2
9.	-12-3=-15	-15	53	-9	45	9	1	15	1	-
25.	-6-4=-10	-10	60	-2	39	2	-	10	1	-
41.	-13-4=-17	-17	65	-9	31	9	1	17	-	3
57.	-9-7=-16	-16	62	-2	35	2	-	16	2	2
13.	-3-+12=-15	-15	68	9	27	-9	3	15	2	-
29.	-2-+11=-13	-13	67	9	24	-9	4	13	3	2
45.	-7-+9=-16	-16	73	2	23	-2	1	16	2	1
61.	-4-+6=-10	-10	73	2	22	-2	2	10	2	1
1.	-6-(-8)=2	2	69	-14	12	14	13	-2	11	5
17.	-1-(-10)=9	9	70	-11	8	11	4	-9	16	2
33.	-3-(-5)=2	2	76	-8	9	8	5	-2	9	1
49.	-2-(-11)=9	9	71	-13	9	13	5	-9	12	3
8.	-16-(-7)=-9	-9	82	-23	9	23	-	9	3	6
24.	-12-(-10)=-2	-2	83	-22	12	22	-	2	3	2
40.	-11-(-2)=-9	-9	83	-13	12	13	2	9	2	1
56.	-7-(-5)=-2	-2	83	-12	10	12	2	2	3	2
7.	8-6=2	2	89	14	6	-14	-	-2	5	-
23.	7-5=2	2	83	12	6	-12	2	-2	8	1
39.	4-2=2	2	86	6	7	-6	1	-2	6	-
55.	9-7=2	2	82	16	9	-16	1	-2	7	1
16.	2-11=-9	-9	65	13	10	-13	5	9	18	2
32.	5-14=-9	-9	72	19	5	-19	5	9	15	3
48.	7-16=-9	-9	66	23	-	-23	3	9	19	12
64.	4-13=-9	-9	81	17	6	-17	2	9	17	4
6.	6+4=10	10	99	2	1	-2	-	-10	-	-
22.	10+8=18	18	97	2	2	-2	1	-18	-	-
38.	2+11=13	13	97	-9	1	9	1	-13	-	1
54.	4+13=17	17	98	-9	1	9	1	-17	-	-
10.	-14+-5=-19	-19	87	-9	9	9	-	19	2	1
26.	-10+-1=-11	-11	87	-9	9	9	-	11	3	-
42.	-7+-5=-12	-12	92	-2	2	2	-	12	2	4
58.	-10+-8=-18	-18	93	-2	4	2	-	18	1	2
14.	-5+-7=-12	-12	87	2	3	-2	5	12	2	3
30.	-6+-8=-14	-14	94	2	2	-2	1	14	1	2
46.	-2+-11=-13	-13	87	9	2	-9	5	13	4	2
62.	-3+-12=-15	-15	92	9	2	-9	4	15	1	-
5.	-3+12=9	9	88	-15	2	15	4	-9	4	2
21.	-1+10=9	9	89	-11	4	11	5	-9	2	-
37.	-4+13=9	9	89	-17	3	17	2	-9	2	4
53.	-2+11=9	9	93	-13	2	13	3	-9	2	-
15.	-6+4=-2	-2	95	-10	4	10	-	2	1	-
31.	-5+3=-2	-2	94	-8	4	8	-	2	1	1
47.	-4+2=-2	-2	93	-6	4	6	-	2	3	-
63.	-8+6=-2	-2	93	-14	5	14	1	2	-	1
3.	12+-3=9	9	91	15	6	-15	1	-9	1	1
19.	7+-5=2	2	91	12	5	-12	1	-2	3	-
35.	15+-6=9	9	90	21	-	-21		-9	1	9
51.	4+-2=2	2	92	6	2	-6	-	-2	5	1
11.	3+-5=-2	-2	93	8	2	-8	3	2	2	-
27.	2+-11=-9	-9	90	13	3	-13	2	9	3	2
43.	6+-8=-2	-2	91	14	1	-14	3	2	3	2
59.	1+-10=-9	-9	91	11	2	-11	7	9	-	1

<u>Error Types:</u>

1. Error Codes for Subtraction:
   a. E2 = Correct Only Operation of Sign
   b. E3 = Changes Operation Sign and Signs of Both Numbers
   c. E4 = Changes Operation Sign and Sign of First Number
2. Error Codes for Addition
   a. E2 = Changes Sign of Second Number
   b. E3 = Changes Sign of First Number
   c. E4 = Changes Sign of Both Numbers

Appendix 3

A Typology of Alternative Algorithms Used by Students in Solving Signed Number Problems

Task Number	Task Type	Alternative Algorithms* 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
4	1-(-10)=11	-9 2	11 1	11 1	-9 2	9 3	-9 2	9 3	-11 1	-9 2	9 3	9 3	-9 2		-9 2	-9 2	-9 2
12	9-(-7)=16	2 2	16 1	16 1	-2 3	2 2	2 2	2 2	16 1	2 1	-2 3	-2 3	2 2		2 2	2 2	2 2
2	-7-9=-16	-2 3	-16 1	16 4	-2 3	-2 3	2 2	2 2	2 2	-16 1	-2 3	-2 3	-16 1		2 2	2 2	-2 3
9	-12-3=-15	-9 2	-15 1	-15 1	-9 2	-9 2	-9 2	-9 2	-9 2	-15 1	9 3	9 3	-15 1		-9 2	-9 2	-9 2
13	-3-(+12)=-15	9 2	-15 1	-15 1	-9 3	-9 3	9 2	9 2	9 2	9 2	-9 3	-9 3	-15 1		9 2	9 2	9 2
1	-6-(-8)=2	-2 4	14 3	14 3	14 3	14 3	-14 2	2 1	2 1	-2 4	-14 2	14 3	-14 2		2 1	-14 2	2 1
8	-16-(-7)=-9	-9 1	-23 2	-23 2	23 3	23 3	-23 2	-9 1	-9 1	-9 1	-23 2	23 3	-23 2		-9 1	-23 2	9 4
7	8-6=2	2 1	14 2	2 1	14 2	14 2	14 2	2 1	2 1	2 1	2 1	-14 3	2 1		2 1	2 1	2 1
16	2-11=-9	-9 1	-13 3	-9 1	13 2	13 2	13 2	9 4	9 4	9 4	-9 1	-13 3	-9 1		9 4	9 4	-9 1
6	6+4=10	2 2	10 1	2 2	10 1	2 2								2 2	10 1	2 2	
10	-14+(-5)=-19	-9 2	-19 1	-9 2	19 4	-9 2								-9 2	-19 1	-9 2	
14	-5+(-7)=-12	-2 3	-12 1	-2 3	12 4	-2 3								2 2	-12 1	-2 3	
5	-3+12=9	9 1	15 3	9 1	-9 4	15 3								-15 2	15 3	9 1	
15	-6+4=-2	-2 1	-10 2	-2 1	-2 1	10 3								-10 2	-10 2	-2 1	
3	12+(-3)=9	9 1	15 2	9 1	-9 4	15 2								15 2	9 1	9 1	
11	3+(-5)=-2	-2 1	-8 3	-2 1	-2 1	8 2								8 2	-8 3	-2 1	

*For an interpretation of each algorithm, refer to the list in Chapter 4.
**The first column in each algorithm gives the row answer to the problem presented under the heading "Task Type." The second column in each algorithm gives the code for the answer given in the first column. Following is the key to the codes:

Answer Codes--

1. Error Codes for Subtraction
   a. 1 = Correct Answer
   b. 2 = Changes Only Operation of Sign
   c. 3 = Changes Operation Sign and Signs of Both Numbers
   d. 4 = Changes Operation Sign and Sign of First Number
2. Error Codes for Addition
   a. 2 = Changes Sign of Second Number
   b. 3 = Changes Sign of First Number
   c. 4 = Changes Sign of Both Numbers

Appendix 3(a)

Wrong Algorithms That Yield Correct Answers
(Only Records That Were Adjusted)

No. of SS		Skill No. 4	12	2	9	13	1	8	7	16	6	10	14	5	15	3	11
2	B	1	1	0	1	1	3	2	1	3	1	1	1	3	2	2	3
	A	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
1	B	0	0	2	2	2	1	1	1	4	1	0	2	1	1	1	1
	A	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
1	B	2	2	2	2	2	4	1	1	4	1	2	3	1	1	1	1
	A	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
1	B	1	1	4	1	1	3	2	1	0	1	2	3	1	1	1	1
	A	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0
1	B	2	3	3	2	3	3	0	2	2	1	4	4	4	1	4	1
	A	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
1	B	2	2	2	2	2	1	1	1	4	1	0	1	3	2	1	3
	A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	B	2	0	2	1	2	4	1	1	4	1	2	3	0	1	1	0
	A	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
1	B	2	2	2	2	0	2	2	1	4	1	2	0	1	1	1	1
	A	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0
1	B	1	1	2	2	2	1	1	1	4							
	A	0	0	0	0	0	0	0	1	0							
3	B	0	2	2	2	2	1	1	1	0							
	A	0	0	0	0	0	0	0	1	0							
4	B	2	2	2	2	2	4	1	1	4							
	A	0	0	0	0	0	0	0	0	0							
2	B	3	1	1	1	1	1	1	1	4							
	A	0	0	1	1	1	1	1	1	0							
1	B	2	2	0	2	2	1	4	0	1							
	A	0	0	0	0	0	0	0	0	0							
1	B	4	4	0	0	1	4	4	0	1							
	A	0	0	0	0	0	0	0	0	0							
1	B	2	2	1	1	2	2	2	0	1							
	A	0	0	0	0	0	0	0	0	0							
1	B	0	3	2	2	2	4	1	1	4							
	A	0	0	0	0	0	0	0	0	0							
1	B	1	1	1	2	1	1	1	4	1							
	A	0	0	0	0	0	0	0	1	0							
2	B	1	1	1	2	1	1	1	4	1							
	A	1	1	0	0	1	1	1	0	0							
1	B	1	2	0	0	1	1	1	1	1							
	A	0	0	0	0	1	1	1	1	0							
1	B	2	2	1	1	0	1	2	0	0							
	A	0	0	0	0	0	0	0	0	0							
1	B	3	2	2	2	2	1	1	1	4							
	A	0	0	0	0	0	0	0	1	0							

B = Before Adjustments
A = After Adjustments

Appendix 4

Correlations, Means and Standard Deviations Before and After Rescoring Items Based on the Underlying Algorithm

Correlation Matrix Among the Sixteen Test Tasks (N = 125)

Task Type	Task No.	Subtraction 4	12	2	9	13	1	8	7	Addition 16	6	10	14	5	15	3	11	Task No.	Before $\bar{x}$	Before S.D.	After $\bar{x}$	After S.D.
L-(-H)	4		.823	.567	.520	.697	.489	.451	.318	.513	.122	.303	.218	.091	-.034	-.032	.153	4	.648	.480	.600	.492
H-(-L)	12	.883		.534	.557	.718	.541	.494	.354	.494	.131	.332	.241	.051	-.018	-.013	.177	12	.680	.468	.624	.486
-L-H	2	.681	.636		.622	.630	.435	.334	.255	.448	.106	.253	.244	.110	.147	-.069	.050	2	.584	.495	.536	.501
-H-L	9	.616	.636	.711		.619	.461	.284	.340	.330	.105	.305	.239	-.014	.002	-.014	.045	9	.576	.496	.536	.501
-L-+H	13	.754	.796	.724	.689		.614	.507	.404	.571	.144	.372	.274	.079	.081	-.053	.144	13	.720	.450	.688	.465
-L-(-H)	1	.733	.777	.629	.697	.910		.643	.438	.583	-.053	.206	.296	.232	.017	.165	.165	1	.744	.438	.664	.474
-H-(-L)	8	.774	.815	.675	.675	.944	.892		.648	.437	.194	.153	.137	.096	.162	.173	.096	8	.824	.382	.696	.462
H-L	7	.461	.497	.511	.551	.633	.595	.604		.283	.219	-.047	.079	.047	.098	.131	-.037	7	.856	.353	.792	.408
L-H	16	.663	.673	.557	.523	.733	.752	.751	.572		-.058	.293	.332	.261	.147	.067	.132	16	.704	.458	.648	.480
H+L	6	.250	.263	.219	.219	.303	.287	.309	.197	.277		-.028	-.023	-.026	.369	-.026	-.026	6	.992	.089	.960	.197
-H+-L	10	.383	.405	.331	.382	.473	.446	.482	.131	.376	.316		.611	.221	.047	.117	.221	10	.912	.284	.888	.317
-L+-H	14	.344	.364	.296	.350	.425	.401	.434	.301	.385	.349	.745		.164	-.064	.043	.043	14	.936	.246	.904	.296
-L+H	5	.280	.248	.229	.178	.308	.338	.317	.193	.323	.316	.598	.659		.441	.348	.565	5	.920	.272	.888	.317
-H+L	15	.310	.331	.314	.261	.393	.368	.402	.213	.353	.465	.627	.690	.710		.313	.441	15	.944	.231	.896	.306
H+-L	3	.274	.246	.220	.220	.259	.285	.267	.158	.269	.288	.623	.607	.623	.654		.348	3	.920	.272	.812	.335
L+-H	11	.352	.374	.249	.249	.389	.363	.398	.175	.295	.302	.649	.632	.728	.681	.669		11	.920	.272	.880	.326

Above The Diagonal Line: Before Rescoring
Below The Diagonal Line: After Rescoring

Appendix 5

Oblique Factor Matrix for Rescored Task Scores

Task Number	Task Type	Loadings F_I	F_{II}
4	S-(-L)	.897	-.022
12	L-(-S)	.915	.000
2	-S-L	.886	-.047
9	-L-S	.887	-.033
13	-S-+L	.919	.013
1	-S-(-L)	.934	.015
8	-L-(-S)	.905	.050
7	L-S	.690	.035
16	S-L	.841	.018
6	L+S	.067	.531
10	-L+-S	.057	.846
14	-S+-L	.045	.871
5	-S+L	-.034	.906
15	-L+S	.000	.931
3	L+-S	-.088	.894
11	S+-L	-.019	.891

Loadings > .3 are underlined

$r_{F_I F_{II}} = .500$

Appendix 6

Response Time --Means and Standard Deviations for the Group of "Experts"

Total Number = 36 out of 127

Task Type	N	Item No.	$\bar{x}$	S.D.	Item No.	N	$\bar{x}$	S.D.	Item No.	N	$\bar{x}$	S.D.	Item No.	N	$\bar{x}$	S.D.
S-(-L)	32	4	12,466	5,451	20	36	11,770	4,537	36	33	12,331	7,077	52	35	11,319	4,122
L-(-S)	31	12	12,656	5,511	28	34	12,222	7,835	44	35	10,380	5,032	60	35	11,701	4,180
-S-L	30	2	13,983	5,645	18	32	13,307	6,524	34	33	10,104	5,643	50	34	13,718	5,928
-L-S	32	9	12,185	5,095	25	33	10,416	5,577	41	30	12,603	6,107	57	33	12,232	5,379
-S-+L	31	13	11,793	6,604	29	32	10,056	3,539	45	35	10,914	4,530	61	35	10,391	5,521
-S-(-L)	33	1	18,975	7,292	17	31	12,291	4,917	33	33	11,388	6,205	49	35	14,687	7,230
-L-(-S)	32	8	17,060	7,667	24	36	10,762	5,028	40	33	11,864	4,541	56	36	10,502	6,653
L-S	35	7	9,918	5,145	23	36	8,767	4,568	39	32	9,492	5,850	55	36	9,844	5,035
S-L	34	16	12,938	6,048	32	35	13,518	6,368	48	35	13,323	6,941	64	32	14,299	7,889
L+S	32	6	7,488	2,791	22	35	6,952	1,810	38	36	8,691	4,536	54	36	7,879	3,180
-L+-S	33	10	10,698	5,566	26	33	7,680	2,517	42	36	9,607	4,828	58	36	8,880	4,267
-S+-L	33	14	10,504	4,773	30	35	10,559	4,614	46	35	11,235	8,488	62	35	9,325	3,531
-S+L	35	5	12,257	5,133	21	35	8,483	2,547	37	32	10,879	4,163	53	35	8,868	3,638
-L+S	36	15	8,386	3,128	31	35	8,818	5,027	47	35	9,728	6,178	63	36	8,720	5,324
L+-S	33	3	13,329	5,198	19	34	8,797	3,852	35	33	13,426	7,402	51	36	8,684	3,721
S+-L	33	11	7,369	2,311	27	35	11,361	5,365	43	35	8,588	4,041	59	34	10,069	4,718

REFERENCES

Bejar, I.I., Weiss, D.J., & Kingsbury, G.G. Calibration of an item pool for the adaptive measurement of achievement. (Research Report 77-5). Minneapolis: Universtiy of Minnesota, Department Psychology, Psychometric Methods Program, September 1977. (NTIS No. AD A04 4828).

Bejar, I.I. & Weiss, D.J. A construct validation of adaptive achievement testing. (Research Report 78-4). Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program, November 1978.

Brown, J.S. & Burton, R.R. Diagnostic models for procedural bugs in basic mathematical skills. Cognitive Science, 1978, 2, 155-192.

Brown, J.S. & Weiss, D.J. An adaptive testing strategy for achievement test batteries. (Research Report 77-6). Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program, October 1977. (NTIS No. AD A046062).

Cazden, C.B. Implications for instructional research in: David Klahr (ed.), Cognition and Instruction. Hillsdale, New Jersey: LEA, 1976

Cronbach, L.J. Coefficient alpha and the internal structure of tests. Psychometrika, 1951, 16, 297-334.

English, R.A., Reckase, M.D., & Patience, W.M. Application of tailored testing to achievement measurement. Behavior Research Methods and Instrumentation, 1977, 9, 158-161.

Glaser, R. Components of a psychology of instruction: Toward a science of design. Review of Educational Research, 1976, 46 (1), 1-24.

Groen, G.J. & Perkum, J.M. A chronometric analysis of simple addition. Psychological Review, 1972, 79, 329-343.

Klahr, D. (Ed.). Cognition and instruction. Hillsdale, New Jersey: LEA, 1976.

Kruskal, J.B., Young, F.W. & Seery, J.B. KYST--A computer program to do multidimensional scaling and unfolding. New Jersey: Bell Laboratories, New Jersey, 1973.

Lachman, R., Lachman, J. & Butterfield, E.C. Cognitive psychology and information processing: An introduction. Hillsdale, New Jersey: LEA, 1979.

Lord, F.M., and Novick, M.R. Statistical theories of mental test scores. Reading: Adison Wesley, 1968.

McKinley, R.L. & Reckase, M.D. A successful application of latent-trait theory to tailored achievement testing. (Research Report 80-1). Columbia: University of Missouri, Department of Educational Psychology, Tailored Testing Research Laboratory, February 1980.

Nie, N.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K.S. & Bent, D.H. Statistical package for the social science (SPSS). New York: McGraw-Hill, 1975.

Resnick, L.B. Task analysis in instructional design: Some cases from Mathematics. In David Klahr (Ed.): Cognition and Instruction. Hillsdale, New Jersey: LEA, 1976.

Tatsuoka, K., Birenbaum, M., Tatsuoka, M. & Baillie, R. A psychometric approach to error-analysis on response patterns. (Research Report 80-2). Urbana, Illinois: University of Illinois, Computer-based Education Research Laboratory, 1980.

Tatsuoka, K.K. & Tatsuoka, M.M. A model for incorporating response-time data in scoring achievement tests. (CERL Report E-7). Final Report for ONR Project No. N00014-78-C-0159. Urbana, Illinois: University of Illinois, Computer-based Education Laboratory, July 1979.

Woods, S.S., Resnick, L.B. & Groen, G.J. An experimental test of five process models for subtraction. Journal of Educational Psychology, 1975, 67, 17-21.

Navy

1 Dr. Jack R. Borsting
Provost & Academic Dean
U.S. Naval Postgraduate School
Monterey, CA 93940

1 Dr. Robert Breaux
Code N-711
NAVTRAEQUIPCEN
Orlando, FL 32813

1 Chief of Naval Education and Training
Liason Office
Air Force Human Resource Laboratory
Flying Training Division
WILLIAMS AFB, AZ 85224

1 Dr. Richard Elster
Department of Administrative Sciences
Naval Postgraduate School
Monterey, CA 93940

1 DR. PAT FEDERICO
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA 92152

1 Mr. Paul Foley
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. John Ford
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Henry M. Halff
Department of Psychology,C-009
University of California at San Diego
La Jolla, CA 92093

1 Dr. Pa k R. Harrison
Psychol. Course Director
LEADERSHI LAW DEPT. (7b)
DIV. OF PROFESSIONAL DEVELOPMMENT
U.S. NAVAL ACADEMY
ANNAPOLIS, MD 21402

1 Psychologist
ONR Branch Office
Bldg 114, Section D
666 Summer Street
Boston, MA 02210

1 Psychologist
ONR Branch Office
536 S. Clark Street
Chicago, IL 60605

Navy

1 CDR Robert S. Kennedy
Head, Human Performance Sciences
Naval Aerospace Medical Research Lab
Box 29407
New Orleans, LA 70189

1 Dr. Norman J. Kerr
Chief of Naval Technical Training
Naval Air Station Memphis (75)
Millington, TN 38054

1 Dr. William L. Maloy
Principal Civilian Advisor for
Education and Training
Naval Training Command, Code 00A
Pensacola, FL 32508

1 Dr. Kneale Marshall
Scientific Advisor to DCNO(MPT)
OP01T
Washington DC 20370

1 CAPT Richard L. Martin, USN
Prospective Commanding Officer
USS Carl Vinson (CVN-70)
Newport News Shipbuilding and Drydock Co
Newport News, VA 23607

1 Dr. James McBride
Navy Personnel R&D Center
San Diego, CA 92152

1 Ted M. I. Yellen
Technical Information Office, Code 201
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA 92152

1 Library, Code P201L
Navy Personnel R&D Center
San Diego, CA 92152

6 Commanding Officer
Naval Research Laboratory
Code 2627
Washington, DC 20390

1 Office of Naval Research
Code 437
800 N. Quincy SStreet
Arlington, VA 22217

5 Personnel & Training Research Programs
(Code 458)
Office of Naval Research
Arlington, VA 22217

Navy

1 Psychologist
ONR Branch Office
1030 East Green Street
Pasadena, CA 91101

1 Office of the Chief of Naval Operations
Research Development & Studies Branch
(OP-115)
Washington, DC 20350

1 LT Frank C. Petho, MSC, USN (Ph.D)
Code L51
Naval Aerospace Medical Research Laborat
Pensacola, FL 32508

1 Dr. Bernard Rimland (03B)
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Worth Scanland
Chief of Naval Education and Training
Code N-5
NAS, Pensacola, FL 32508

1 Dr. Robert G. Smith
Office of Chief of Naval Operations
OP-987H
Washington, DC 20350

1 Dr. Alfred F. Smode
Training Analysis & Evaluation Group
(TAEG)
Dept. of the Navy
Orlando, FL 32813

1 Dr. Richard Sorensen
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Ronald Weitzman
Code 54 WZ
Department of Administrative Sciences
U. S. Naval Postgraduate School
Monterey, CA 93940

1 Dr. Robert Wisher
Code 309
Navy Personnel R&D Center
San Diego, CA 92152

1 DR. MARTIN F. WISKOFF
NAVY PERSONNEL R& D CENTER
SAN DIEGO CA 92152

Army

1 Technical Director
U. S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Myron Fischl
U.S. Army Research Institute for the
Social and Behavioral Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Dexter Fletcher
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria,VA 22333

1 Dr. Michael Kaplan
U.S. ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

1 Dr. Milton S. Katz
Training Technical Area
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Harold F. O'Neil, Jr.
Attn: PERI-OK
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Mr. Robert Ross
U.S. Army Research Institute for the
Social and Behavioral Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Robert Sasmor
U. S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Commandant
US Army Institute of Administration
Attn: Dr. Sherrill
FT Benjamin Harrison, IN 46256

1 Dr. Frederick Steinheiser
U. S. Army Reserch Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Army

1 Dr. Joseph Ward
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Air Force

1 Air Force Human Resources Lab
AFHRL/MPD
Brooks AFB, TX 78235

1 Dr. Earl A. Alluisi
HQ, AFHRL (AFSC)
Brooks AFB, TX 78235

1 Research and Measurment Division
Research Branch, AFMPC/MPCYPR
Randolph AFB, TX 78148

1 Dr. Malcolm Ree
AFHRL/MP
Brooks AFB, TX 78235

1 Dr. Marty Rockway
Technical Director
AFHRL(OT)
Williams AFB, AZ 58224

Civil Govt

1 Dr. Andrew R. Molnar
Science Education Dev.
and Research
National Science Foundation
Washington, DC 20550

1 Dr. Vern W. Urry
Personnel R&D Center
Office of Personnel Management
1900 E Street NW
Washington, DC 20415

1 Dr. Joseph L. Young, Director
Memory & Cognitive Processes
National Science Foundation
Washington, DC 20550

Marines

1 H. William Greenup
Education Advisor (E031)
Education Center, MCDEC
Quantico, VA 22134

1 Director, Office of Manpower Utilization
HQ, Marine Corps (MPU)
BCB, Bldg. 2009
Quantico, VA 22134

1 DR. A.L. SLAFKOSKY
SCIENTIFIC ADVISOR (CODE RD-1)
HQ, U.S. MARINE CORPS
WASHINGTON, DC 20380

Other DoD

12 Defense Technical Information Center
Cameron Station Bldg 5
Alexandria, VA ,14
Attn: TC

1 Dr. William Graham
Testing Directorate
MEPCOM/MEPCT-P
Ft. Sheridan, IL 60037

1 Military Assistant for Training and
Personnel Technology
Office of the Under Secretary of Defense
for Research & Engineering
Room 3D129, The Pentagon
Washington, DC 20301

1 MAJOR Wayne Sellman, USAF
Office of the Assistant Secretary
of Defense (MRA&L)
3B930 The Pentagon
Washington, DC 20301

1 DARPA
1400 Wilson Blvd.
Arlington, VA 22209

CoastGuard

1 Mr. Thomas A. Warm
U. S. Coast Guard Institute
P. O. Substation 18
Oklahoma City, OK 73169

Non Govt

1 Charles Myers Library
Livingstone House
Livingstone Road
Stratford
London E15 2LJ
ENGLAND

1 Dr. Kenneth E. Clark
College of Arts & Sciences
University of Rochester
River Campus Station
Rochester, NY 14627

1 Dr. Norman Cliff
Dept. of Psychology
Univ. of So. California
University Park
Los Angeles, CA 90007

1 Dr. William E. Coffman
Director, Iowa Testing Programs
334 Lindquist Center
University of Iowa
Iowa City, IA 52242

1 Dr. Meredith P. Crawford
American Psychological Association
1200 17th Street, N.W.
Washington, DC 20036

1 Dr. Leonard Feldt
Lindquist Center for Measurment
University of Iowa
Iowa City, IA 52242

1 Dr. Richard L. Ferguson
The American College Testing Program
P.O. Box 168
Iowa City, IA 52240

1 Dr. Victor Fields
Dept. of Psychology
Montgomery College
Rockville, MD 20850

1 Univ. Prof. Dr. Gerhard Fischer
Liebiggasse 5/3
A 1010 Vienna
AUSTRIA

1 DR. ROBERT GLASER
LRDC
UNIVERSITY OF PITTSBURGH
3939 O'HARA STREET
PITTSBURGH, PA 15213

1 Dr. John R. Frederiksen
Bolt Beranek & Newman
50 Moulton Street
Cambridge, MA 02138

1 Dr. Erling B. Andersen
Department of Statistics
Studiestraede 6
1455 Copenhagen
DENMARK

1 1 psychological research unit
Dept. of Defense (Army Office)
Campbell Park Offices
Canberra ACT 2600, Australia

1 Dr. Isaac Bejar
Educational Testing Service
Princeton, NJ 08450

1 Dr. Werner Birke
DezWPs im Streitkraefteamt
Postfach 20 50 03
D-5300 Bonn 2
WEST GERMANY

1 Dr. Nicholas A. Bond
Dept. of Psychology
Sacramento State College
600 Jay Street
Sacramento, CA 95819

1 Dr. Robert Brennan
American College Testing Programs
P. O. Box 168
Iowa City, IA 52240

1 DR. C. VICTOR BUNDERSON
WICAT INC.
UNIVERSITY PLAZA, SUITE 10
1160 SO. STATE ST.
OREM, UT 84057

1 Dr. John B. Carroll
Psychometric Lab
Univ. of No. Carolina
Davie Hall 013A
Chapel Hill, NC 27514

1 Library
HumRRO/Western Division
27857 Berwick Drive
Carmel, CA 93921

1 Dr. Steven Hunka
Department of Education
University of Alberta
Edmonton, Alberta
CANADA

Non Govt

1 Dr. Earl Hunt
Dept. of Psychology
University of Washington
Seattle, WA 98105

1 Dr. Huynh Huynh
College of Education
University of South Carolina
Columbia, SC 29208

1 Professor John A. Keats
University of Newcastle
AUSTRALIA 2308

1 Mr. Marlin Kroger
1117 Via Goleta
Palos Verdes Estates, CA 90274

1 Dr. Charles Lewis
Faculteit Sociale Wetenschappen
Rijksuniversiteit Groningen
Oude Boteringestraat
Groningen
NETHERLANDS

1 Dr. Robert Linn
College of Education
University of Illinois
Urbana, IL 61801

1 Dr. Frederick M. Lord
Educational Testing Service
Princeton, NJ 08540

1 Dr. Gary Marco
Educational Testing Service
Princeton, NJ 08450

1 Dr. Scott Maxwell
Department of Psychology
University of Houston
Houston, TX 77004

1 Dr. Samuel T. Mayo
Loyola University of Chicago
820 North Michigan Avenue
Chicago, IL 60611

1 Dr. Ron Hambleton
School of Education
University of Massechusetts
Amherst, MA 01002

Non Govt

1 Dr. Melvin R. Novick
356 Lindquist Center for Measurment
University of Iowa
Iowa City, IA 52242

1 Dr. Jesse Orlansky
Institute for Defense Analyses
400 Army Navy Drive
Arlington, VA 22202

1 Dr. James A. Paulson
Portland State University
P.O. Box 751
Portland, OR 97207

1 MR. LUIGI PETRULLO
2431 N. EDGEWOOD STREET
ARLINGTON, VA 22207

1 DR. DIANE M. RAMSEY-KLEE
R-K RESEARCH & SYSTEM DESIGN
3947 RIDGEMONT DRIVE
MALIBU, CA 90265

1 MINRAT M. L. RAUCH
P II 4
BUNDESMINISTERIUM DER VERTEIDIGUNG
POSTFACH 1328
D-53 BONN 1, GERMANY

1 Dr. Mark D. Reckase
Educational Psychology Dept.
University of Missouri-Columbia
4 Hill Hall
Columbia, MO 65211

1 Dr. Andrew M. Rose
American Institutes for Research
1055 Thomas Jefferson St. NW
Washington, DC 20007

1 Dr. Leonard L. Rosenbaum, Chairman
Department of Psychology
Montgomery College
Rockville, MD 20850

1 Dr. Chester Harris
School of Education
University of California
Santa Barbara, CA 93106

1 Dr. Lloyd Humphreys
Department of Psychology
University of Illinois
Champaign, IL 61820

Non Govt

1 Dr. Ernst Z. Rothkopf
Bell Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974

1 Dr. Lawrence Rudner
403 Elm Avenue
Takoma Park, MD 20012

1 Dr. J. Ryan
Department of Education
University of South Carolina
Columbia, SC 29208

1 PROF. FUMIKO SAMEJIMA
DEPT. OF PSYCHOLOGY
UNIVERSITY OF TENNESSEE
KNOXVILLE, TN 37916

1 DR. ROBERT J. SEIDEL
INSTRUCTIONAL TECHNOLOGY GROUP
HUMRRO
300 N. WASHINGTON ST.
ALEXANDRIA, VA 22314

1 Dr. Kazuo Shigemasu
University of Tohoku
Department of Educational Psychology
Kawauchi, Sendai 980
JAPAN

1 Dr. Edwin Shirkey
Department of Psychology
University of Central Florida
Orlando, FL 32816

1 Dr. Robert Smith
Department of Computer Science
Rutgers University
New Brunswick, NJ 08903

1 Dr. Richard Snow
School of Education
Stanford University
Stanford, CA 94305

1 Professor Donald Fitzgerald
University of New England
Armidale, New South Wales 2351
AUSTRALIA

1 Dr. Edwin A. Fleishman
Advanced Research Resources Organ.
Suite 900
4330 East West Highway
Washington, DC 20014

Non Govt

1 Dr. Robert Sternberg
Dept. of Psychology
Yale University
Box 11A, Yale Station
New Haven, CT 06520

1 DR. PATRICK SUPPES
INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES
STANFORD UNIVERSITY
STANFORD, CA 94305

1 Dr. Hariharan Swaminathan
Laboratory of Psychometric and Evaluation Research
School of Education
University of Massachusetts
Amherst, MA 01003

1 Dr. Brad Sympson
Psychometric Research Group
Educational Testing Service
Princeton, NJ 08541

1 Dr. David Thissen
Department of Psychology
University of Kansas
Lawrence, KS 66044

1 Dr. Robert Tsutakawa
Department of Statistics
University of Missouri
Columbia, MO 65201

1 Dr. J. Uhlaner
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

1 Dr. Howard Wainer
Bureau of Social SCience Research
1990 M Street, N. W.
Washington, DC 20036

1 Dr. Phyllis Weaver
Graduate School of Education
Harvard University
200 Larsen Hall, Appian Way
Cambridge, MA 02138

1 DR. SUSAN E. WHITELY
PSYCHOLOGY DEPARTMENT
UNIVERSITY OF KANSAS
LAWRENCE, KANSAS 66044

1 Wolfgang Wildgrube
Streitkraefteamt
Box 20 50 03
[illegible]

DATE

ILMED

-8